

DÆDALUS

JOURNAL OF THE AMERICAN ACADEMY OF ARTS AND SCIENCES
FALL 1994

US: \$7.95

CANADA: \$10.35

HEALTH



&

WEALTH



Max S. Cynader • Robert G. Evans • John W. Frank
Clyde Hertzman • Naoki Ikegami • John N. Lavis
Margaret Lock • Michael G. Marmot • Naoko T. Miyaji
J. Fraser Mustard • Paul M. Romer • Nathan Rosenberg
Gregory L. Stoddart • S. Leonard Syme • Richard G. Wilkinson
Michael C. Wolfson

PRODUCED BY UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

BOARD OF EDITORS

David Apter, David Baltimore, Daniel Bell, Isaiah Berlin,
Guido Calabresi, Natalie Z. Davis, Wendy Doniger,
Clifford Geertz, Stephen J. Greenblatt, Vartan Gregorian,
Stanley Hoffmann, Gerald Holton, Donald Kennedy,
Sally F. Moore, Gian-Carlo Rota, W. G. Runciman,
Amartya K. Sen, Steven Weinberg

STEPHEN R. GRAUBARD
Editor of the Academy and of Dædalus

PHYLLIS S. BENDELL
Associate Editor

AMY E. TAYLOR
Assistant Editor

JENNIFER A. POOLE
Editorial and Circulation Office Assistant

Cover design by Michael Schubert, Director of Ruder-Finn Design



*Printed on
recycled paper*

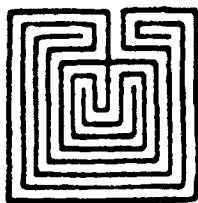
DÆDALUS

JOURNAL OF THE AMERICAN ACADEMY OF ARTS AND SCIENCES

Health and Wealth

Fall 1994

Issued as Volume 123, Number 4, of the
Proceedings of the American Academy of Arts and Sciences



Fall 1994, "Health and Wealth."

Issued as Volume 123, Number 4, of the Proceedings of the American Academy of Arts and Sciences.

© 1994 by the American Academy of Arts and Sciences,
136 Irving Street, Cambridge, MA 02138.

Library of Congress Catalog Number 12-30299.

Editorial Offices: Dædalus, Norton's Woods,
136 Irving Street, Cambridge, MA 02138.

Dædalus (ISSN 0011-5266) is published quarterly by the American Academy of Arts and Sciences. Subscription rates: for individuals—\$30, one year; \$55, two years; \$75, three years; for institutions—\$45, one year; \$75, two years; \$100, three years. Foreign subscribers must add \$7.00 per year to the price of subscription. Replacement copies for damaged or misrouted issues will be sent free of charge up to six months from the date of original publication. Thereafter, back copies are available for the current cover price plus postage and handling.

All subscription orders, single-copy orders, and change-of-address information must be sent in writing to the Dædalus Business Office, 136 Irving Street, Suite 100, Cambridge, MA 02138, U.S.A. (Telephone: 617-491-2600).

Printed in the United States of America.

Printing Office: 2901 Byrdhill Road, Richmond, VA 23228.

U.S.A. newsstand distribution by Eastern News Distributors, Inc.,
2020 Superior Street, Sandusky, OH 44870 (Telephone: 800-221-3148).

Second-class postage paid at Boston, MA, and at additional mailing offices.

Postmaster: Send address changes to DÆDALUS,
136 Irving Street, Suite 100, Cambridge, MA 02138, U.S.A.



Contents

- V Preface
John W. Frank and J. Fraser Mustard
- 1 The Determinants of Health from a Historical Perspective
Robert G. Evans
- 21 Health Care as a Threat to Health: Defense, Opulence,
and the Social Environment
John N. Lavis and Gregory L. Stoddart
- 43 Can We Have Too Much Health Care?
Richard G. Wilkinson
- 61 The Epidemiological Transition: From Material Scarcity
to Social Disadvantage?
S. Leonard Syme
- 79 The Social Environment and Health
Naoko T. Miyaji and Margaret Lock
- 87 Monitoring Motherhood: Sociocultural and Historical
Aspects of Maternal and Child Health in Japan
Naoki Ikegami
- 113 Efficiency and Effectiveness in Health Care
Nathan Rosenberg
- 127 How the Developed Countries Became Rich
Paul M. Romer
- 141 Economic Growth and Investment in Children
Max S. Cynader
- 155 Mechanisms of Brain Development and Their Role in
Health and Well-Being
Clyde Hertzman
- 167 The Lifelong Impact of Childhood Experiences: A
Population Health Perspective

- Michael C. Wolfson**
181 **Toward a System of Health Statistics**
- Michael G. Marmot**
197 **Social Differentials in Health Within and Between Populations**

ADVISORY BOARD

J. Fraser Mustard, Toshiyuki Furukawa, Robert G. Evans, John W. Frank, Toshio Sasaki, Kazuhiko Atsumi, Kei Takeuchi, Max S. Cynader, Clyde Hertzman, Naoki Ikegami, John N. Lavis, Margaret Lock, Michael G. Marmot, Naoko T. Miyaji, Kanehisa Morimoto, Marc Renaud, Paul M. Romer, Nathan Rosenberg, Gregory L. Stoddart, S. Leonard Syme, Richard G. Wilkinson, Michael C. Wolfson

Preface to the Issue “Health and Wealth”

IN THE WINTER OF 1977, *Dædalus* published an issue entitled “Doing Better and Feeling Worse: Health in the United States.” Inspired largely by the late Dr. John H. Knowles, responding to an initiative launched by Nelson Rockefeller in his Commission on Critical Choices for Americans, the authors were invited to reflect on what a viable health policy for the nation might be. Dr. Knowles expressed sentiments common at the time when he wrote:

The American people have clearly come to expect much from medicine, especially in recent years, but they have matched these rapidly rising expectations with rising anxieties over the cost, quality, and accessibility of health services. Many within the profession believe that both the expectations and the anxieties are the results of a definition of health that has become much too broad and that certain limits on the responsibility to society of the medical profession need to be established. At a time when the relations between health, illness, and medicine are being viewed in wholly new ways and are acquiring new meanings in our culture, some talk of too much “medicalization,” and argue for “demedicalization”—a return to a simpler set of health practices—even as others call for greater efforts toward realizing the advantages that only an improved science and technology can provide. They say that we are only at the beginning

VI *Dædalus*

of a “scientific revolution” in the health field, although, if that is so, it becomes all the more incumbent upon us to develop effective ways of assessing the economic, social, ethical, and medical impact of new technologies before they are generally introduced into practice.¹

Three of the authors, all now deceased, wrote memorably about the problems of health care, as they perceived them. Aaron Wildavsky, showing his customary candor, expressed the views of others when he said:

According to the Great Equation, Medical Care equals Health. But the Great Equation is wrong. More available medical care does not equal better health. The best estimates are that the medical system (doctors, drugs, hospitals) affects about 10 per cent of the usual indices for measuring health: whether you live at all (infant mortality), how well you live (days lost due to sickness), how long you live (adult mortality). The remaining 90 per cent are determined by factors over which doctors have little or no control, from individual life style (smoking, exercise, worry), to social conditions (income, eating habits, physiological inheritance), to the physical environment (air and water quality). Most of the bad things that happen to people are at present beyond the reach of medicine.²

Dr. Lewis Thomas, in his inimitable prose—no American doctor of his day wrote more felicitously—described what medicine could and could not do to improve the health of individuals. With his characteristic ability to be precise in his observations, philosophical in his recommendations, Thomas wrote:

“Seeing the doctor” also includes an overwhelming demand for reassurance. Transient upper-respiratory infections and episodes of gastroenteritis account for most of the calls on a doctor because of illness, and an even greater number of calls are made by people who have nothing at all the matter with them. It is often claimed that these are mostly unhappy individuals, suffering from psychoneuroses, in need of compassionate listening on the part of the physician, but a large number of patients who find themselves in doctors’ offices or hospital clinics will acknowledge themselves to be in entirely good health; they are there because of a previous appointment in connection with an earlier illness, for a “checkup,” or for a laboratory test, or simply for reassurance that they are not coming down with something serious—cancer, or heart disease, or whatever. Or they may have come to the doctor for advice about living: what should their

diet be?, should they take a vacation?, what about a tranquilizer for everyone's inevitable moments of agitation and despair? I know a professor of pediatrics who has received visits from intelligent, well-educated parents who only want to know if their child should start Sunday school.

The system is being overused, swamped by expectant overdemand for services that are frequently trivial or unproductive. The public is not sufficiently informed of the facts about things that medicine can and cannot accomplish. Medicine is surely not in possession of special wisdom about how to live a life.

It needs to be said more often that human beings are fundamentally tough, resilient animals, marvelously made, most of the time capable of getting along quite well on their own. The health-care system should be designed for use when it is really needed and when it has something of genuine value to offer. If designed, or redesigned, in this way, the system would function far more effectively, and would probably cost very much less.³

Dr. Walsh McDermott, with his vast medical experience, in caring for patients, but also in developing treatments for syphilis, typhus and typhoid fevers, bacterial pneumonias, and tuberculosis, wrote soberly as someone who understood what the public health system could do, what the individual physician ought to strive to do, and what technology might in time do. Again, he understood how complex was the problem of evaluating the work of the individual physician, the worth of new technologies and drugs. McDermott wrote:

Health care is to some extent performed by oneself, but it is for the most part performed by others. These "others" are found in one of two systems: the public health system, which embraces the activities of a wide range of health professionals serving with various degrees of autonomy and only loosely and indirectly related to the physician, and the personal-encounter-physician system, in which a considerably narrower range of health professionals works in a much more direct relationship to the individual doctor. The biomedical influence of these two systems affects us all. We are also affected by a third factor—the style of living permitted by our socioeconomic or technological status. The impact of this last and of the public health-care system are closely intertwined, but they can be analyzed separately in a fair number of instances, so that it then becomes possible to relate their functioning to outcome. Thus, for at least one of the two health-

VIII *Dædalus*

care systems, that of public health, we can measure the disease pattern of a community, and therefore its level of health, and rank it in these terms with other communities.

With the personal-encounter-physician system the situation is quite different. Here we are measuring the activities of individual physicians and their various associates rather than the diseases present in a community, a much more difficult undertaking. Everything the physician does is some combination of technology and what we might call "samaritanism." Technology includes all those products of medical science that are useful in preventing or altering disease in a specific fashion. "Samaritanism" is the collection of acts that provide reassurance, or at least support, to someone troubled by disease or illness. These two functions—the technological and the samaritan—are separable in theory but not in practice. A doctor cannot get a passing grade by being proficient at one but not the other; he must be good at both, for they are to be regarded as opposite sides of the same coin.⁴

McDermott's phrase, the "style of living permitted by our socioeconomic or technological status," goes some distance towards explaining the concerns that predominate in the present issue of *Dædalus*, "Health and Wealth." The issue, inspired by the work of the Population Health Program of The Canadian Institute for Advanced Research, seeks to explain how health in the contemporary world is a function of much more than funds made available through systems of private or public health insurance. These are of course important, but so also is the class structure of the particular society, the work that men and women do, the way that children are cared for, the cultural definitions of well-being and health.

All the authors, cognizant of what national health-care systems can provide, not least in guaranteeing adequate medical care to all, know that no single society has developed the perfect health-care system, that it is only for the others to adopt it for them to enjoy equal benefits. Dr. J. Fraser Mustard, for example, though certain that the Canadian health-care system is in many ways superior to the American, is inclined neither to argue nor even to mention this issue in his own introductory essay. His concern is with something of greater importance, as relevant to the United States as it is to Canada, the United Kingdom, Japan, or any other modern industrial society. In the conclusion to the essay he and John W. Frank have written, they say:

It is time to integrate our understanding of the determinants of health and the determinants of economic growth. Governments and their societies are mistaken to concentrate on the economics of business cycles rather than the long-term forces affecting economic growth, prosperity, and health and well-being. Fogel has concluded that 50 percent of the economic growth in Britain since the Industrial Revolution has been due to better nutrition of the population. A society that handicaps large segments of its population during periods of major technological change may be handicapping its future economic growth. We now have a better understanding of the relationships among economic growth, prosperity, and health and well-being, and the need for a long-term integrated perspective on the determinants of health and economic growth. Can we make intelligent and wise use of this understanding?⁵

This concern with economic growth and its relevance for health, echoed also in other articles, is not simply a plea for each country, seeking to improve the health of its citizens, to determine to become more prosperous. Professor Robert Evans, an economist, and another of the principal architects of the study, expresses the motif of the volume in a section that bears the provocative heading "The Determinants of Health: More Ways to Skin a Cat." Evans writes:

The understanding that health, of both individuals and populations, is powerfully influenced by factors other than health care is ancient. It has been somewhat eclipsed, over the past century, by the extraordinary effect of the application of science and technology to medicine. But it is now enjoying a considerable resurgence, not only because of the development of a more measured and critical view of the contributions of health care but also because of the research findings on the determinants of health. Some of these are coming from sociology and psychology, always a bit skeptical of the claims of "mainstream" medicine. But surprisingly supportive evidence is also emerging from epidemiology, molecular biology, neurophysiology, and even primatology.

Possible biological foundations are beginning to emerge for the widely observed correlation between prosperity and health, a correlation which includes but goes far beyond the obvious connection between poverty and deprivation and corresponding physiological damage. There is extensive evidence for a correlation between health and hierarchical position within populations. This may reflect the effects on health of differences in material circumstances, education,

status, or sense of control over one's environment. If the relationships between health and "other things"—prosperity, self-esteem, sense of control—within populations also underlie the differences between populations (either different societies at the same time, or "the same" society at different times), then, quite conceivably, economic growth and development in itself is a powerful contributor to health.

But the connection is not simple. At a very low level of development, economic growth may mean more and better food, housing, clothing—(relatively) well-understood pathways to better health. But in more highly developed societies, the connections may be more subtle. Does prosperity per se contribute to health, or is it a marker for other forms of progress? Economic output is an indicator of what a population can do. But it is not the only one. Ability is a powerful source of self-esteem. But it is not the only one. So the pathways are not clear. But it is clear that such pathways exist. The connections between health and social environment, broadly defined, are very real and may be very powerful.⁶

Other of the authors, writing in a comparable vein, ask such provocative questions as: Is it legitimate to argue that many countries, and not only the United States, are already investing too much in health care, perceived largely as medical care? Is it possible that they are mistaken in allocating such large sums for medicine and medical research, often intended to prolong life—a year or two in many instances—for those who are very aged and chronically ill? Might they not be better advised to invest more, for example, in activities calculated to create economic growth? Would this not produce better health for many millions? Indeed, might a larger investment in day care not bring more significant health benefits than those now regularly made from public appropriations for other more limited purposes? Some authors argue, in effect, that it is the social fabric of society—its economic and social class system, for example—that may be more determining of health and well-being than other factors commonly said to be important. How one goes about changing the social fabric of a society, assuming it is indeed relevant for health, and how one treats the problems created by organizations, and not only those that operate in the workplace, which so heavily emphasize hierarchy, are matters of extraordinary complexity. It cannot be at all surprising that few politicians are inclined to treat such matters. It is much easier to argue about the

costs of specific medical benefits, though even such public discussion will cause many gouty toes to be trod on.

It is no secret that even when societies invest very large sums in studying and treating specific maladies—heart disease is an obvious example—the remedies proposed, though obviously significant given the extraordinary medical and surgical triumphs of this century, are able to do relatively little to arrest the suffering of many who in the end experience such disease. As for the causes of coronary heart disease, while no one would seriously question the very large contributions made by medical research in demonstrating how much diet and smoking, for example, are important factors, each being responsible for much of the coronary heart disease that takes so many lives, they do not, by themselves, or in conjunction with other similar factors, explain the high incidence of heart failure. More needs to be known, and the suggestion that social causes may be more important than those defined as physical—that environmental factors, very broadly defined—may be influential in ways that are too rarely studied is a proposition calculated to raise controversy, not least among those who have become habituated to believing the opposite.

That Japan figures prominently in this volume is in no way surprising. Its infant and adult mortality statistics, like the tables that provide such eloquent testimony on the extraordinary changes in life expectancy that have occurred since the end of World War II, cannot fail to interest those concerned with health and health care in modern industrial societies. Again, and not only because it emphasizes values presumably no longer comparably prized in many countries in the West, a good deal of emphasis is given to the fact that motherhood and family bonds remain central tenets of Japan's cultural and social ideology. That this system imposes restraints, not least on young Japanese women who feel the obligation to abide by the unwritten rules of the society, is acknowledged. Whether such restraints will continue to be accepted, it is of course impossible to say. Still, in a society where long-term and acute care is recognized to be inferior, where hospitals and medical research do not command the resources common in many places in the West, the indices of health, as conventionally given, are exceptionally favorable. Why? Are the Japanese healthier than others? Do our present indices of mortality and morbidity demonstrate that? Or,

XII *Dædalus*

ought other criteria for the health of a single population to be considered in our statistical studies?

If the object of a health-care system is not only to extend life but to make life better, the comparison with conditions that obtain elsewhere must concern all who have any interest in health-care reform. Just as the Japanese “mama” will not soon find its analog in the West, so the care of children, in so many ways dependent on complex cultural and social conditions involving individual and collective preferences, must take account of what political agencies are prepared to do in instituting and sustaining family support policies. What importance is given the health of a large nonvoting population, children who are never in a position to press for specific reforms? The adult population is the decisive force, and in today’s democratic societies, where so many clamor for attention and resources, children are not always given the highest priority. That the medical discoveries of the twentieth century have largely eradicated the fatal childhood diseases of an earlier day, that many more children live to become adults, that infant mortality everywhere in the developed world is on the decline, are generally accepted facts. It is also a fact that vast numbers of children are neglected—the term child abuse has become commonplace—and it is not only in America that cities have become dangerous, for children as for adults. How all this affects a nation’s health, what importance it has for the social environment more generally, are issues that must figure on the new public health agenda.

In the recent health-care debates in the United States, much attention has been given to the costs of proposed changes. If universal access to health care has seemed to be an unattainable goal—not least for fiscal reasons—if there has been as much talk about those who would be “losers” as those who would presumably be “winners,” the rhetoric alone says a good deal about how the issue of a nation’s health is perceived by those who belong to its political class, and also by those professionals who can be presumed to have an economic interest in the matter. The proposition that social and cultural factors need always to be considered in any discussion of health, that sensitive moral and ethical questions are always present, and that they involve much more than simply helping those who are in need, is too rarely acknowledged. Social class and race are virtually forbidden subjects, touching as they do on matters known

to be sensitive, best left untouched by those who would prefer to find a magic organizational solution for issues that are fundamentally cultural and political.

There is no nostalgia in this volume for conditions that once obtained, but also very little celebration of what now exists, either in those countries where all are fully insured, or in those where they are not. This is not to deny the real medical progress that has been made in recent decades. Nor is it to disclaim the notion that certain democratic societies are more solicitous of the health of their populations than are others. It is simply to acknowledge that disease and health are almost certainly caused by factors more complex than those acknowledged by politicians, who have good reasons for wanting to keep costs down, or by others who stand to benefit from keeping things as they are. In a study that looks closely at health conditions and social preoccupations in societies as culturally different as the Canadian, the Japanese, the British, and the American, there are abundant lessons for those prepared to acknowledge that health comes from more than a conquest of disease, or indeed from a provision of sufficient hospital beds, though both are absolutely crucial.

A great debt is owed the Honda Foundation for making this issue possible. It demonstrates the willingness of a Japanese foundation in partnership with institutions in other countries (in this case, The Canadian Institute for Advanced Research) to support the kinds of cross-cultural studies that remain too uncommon even today. It was both a pleasure and an honor to have Mrs. Honda present at the Toronto symposium. The work of Professor Toshiyuki Furukawa and Professor Robert Evans in organizing the scientific part of the meeting must be noted. We owe them our gratitude. Finally, special thanks are due the Canadian Department of External Affairs which, together with the Honda Foundation, provided the funds that made this *Dædalus* issue possible.

The discussions of the draft essays in Toronto led to very spirited exchanges. On the kinds of questions raised in this issue on "Health and Wealth," there can be no universal agreement, but the object of publication in *Dædalus* is precisely to initiate a dialogue that needs to take place. We are very grateful to Dr. J. Fraser Mustard and his colleagues who recognized very early that something other than a verbatim conference transcript was needed if the unconventional

XIV *Dædalus*

views of a good number of participants were to be aired. The authors agreed to make major revisions to their texts, taking into account the frank discussions of men and women who did not always come from their own professions or academic disciplines. Their views, as published in this issue, are intended to provoke debate. We hope they will.

S.R.G.

ENDNOTES

- ¹John H. Knowles, "Introduction," *Dædalus* 106 (1) (Winter 1977): 4–5.
- ²Aaron Wildavsky, "Doing Better and Feeling Worse: The Political Pathology of Health Policy," *Dædalus* 106 (1) (Winter 1977): 105.
- ³Lewis Thomas, "On the Science and Technology of Medicine," *Dædalus* 106 (1) (Winter 1977): 45–46.
- ⁴Walsh McDermott, "Evaluating the Physician and His Technology," *Dædalus* 106 (1) (Winter 1977): 136.
- ⁵John W. Frank and J. Fraser Mustard, "The Determinants of Health from a Historical Perspective," *Dædalus* 123 (4) (Fall 1994): 15.
- ⁶Robert G. Evans, "Health Care as a Threat to Health: Defense, Opulence, and the Social Environment," *Dædalus* 123 (4) (Fall 1994): 33–34.

The Determinants of Health from a Historical Perspective

OUR CONCEPTIONS OF THE DETERMINANTS of health change periodically. At certain times socioeconomic factors have figured predominantly in policy-making; at other times the emphasis has been largely on identifying the causes of disease and treating the sick. Theories about the determinants of health—indeed, the definition of health—necessarily affect how illness is defined, what public policies are initiated, and how resources are allocated.

The National Health Services was introduced in the United Kingdom in 1948. At that time, it was believed that the gradient in health across the social classes (the highest social class had the lowest mortality rates, and the lowest social classes had the highest mortality rates) would be decreased if the financial barriers to health care were removed.¹ Similar arguments were used when Canada introduced its national insurance program for health-care services nearly two decades later. In the 1970s the Merrison Royal Commission on the National Health Service in the United Kingdom² was surprised to find that the gradient in mortality across social classes had actually widened even though mortality rates had continued to fall. In 1977, the Labour government established a research group, headed by Douglas Black, the Chief Scientist in the Department of Health and Social Security, to look more broadly at the factors influencing health. A major conclusion of this compel-

John W. Frank is Director of Research at the Institute for Work and Health, Toronto, and a Fellow of The Canadian Institute for Advanced Research's Population Health Program at the University of Toronto.

J. Fraser Mustard is President of The Canadian Institute for Advanced Research, Toronto, Canada.

ling report, often referred to as the "Black Report,"³ was that, while health care contributed to improved health and well-being, there were socioeconomic factors of equal or greater importance in determining health and well-being. These factors were primarily causing the gradient in health across social classes.

The publication of the "Black Report" unleashed a vigorous and, at times, acrimonious debate and a deluge of studies. Black's working group was itself split over how resources should be allocated to reduce the inequalities in health. As Sir Douglas Black explained:

We were all agreed that education and preventive measures, specifically directed towards the socially deprived, were necessary. But the sociological members of the group (Townsend and Smith) considered that the consequent expenditure should be obtained by diversion from the acute services. On the other hand the medical members—and that means both of us (Black and Morris)—felt that the acute services played a vital part in the prevention of chronic disability and could not be further cut back without serious effects on emergency care, on the training of doctors for both hospital work and for family practice and on the length of waiting lists. We spent a long time, without real success, trying to resolve this matter.⁴

The relative importance of investments in health care versus investments in other determinants of health is still an unresolved issue. The "Black Report," recent population-based epidemiological studies, and new insights from medical science have begun to give us a better understanding of how socioeconomic factors influence the health and well-being of populations.⁵ Some quite striking perspectives have emerged from recent studies of the records in Western countries over the last three hundred years. The early period of our history provides some insights about our changing social environments and health.

PRE-INDUSTRIAL REVOLUTION

Ester Boserup has provided an informative account of hunter-gatherer societies and the Agricultural Revolution.⁶ The hunter-gatherer societies had limited supplies of food and were constantly moving. Since only a small number of young children could be coped with under such conditions, birth spacing, infanticide, and short life expectancy constrained population growth. These groups,

in addition to being affected by changes in food sources, were also exposed to physical hazards, predators, and, presumably, in certain environments, to pathogens.

The basic social structure of the hunter-gatherer societies was the troop or tribe (usually fewer than one hundred individuals). Because the groups were so small, there was a great deal of social interaction and support. Since social support appears to influence health and well-being,⁷ the individuals in these simple social units may not have suffered from the negative effects on health of social deprivation that are found in societies today.

A different social framework emerged in the agricultural evolutionary stage, about ten thousand years ago, and began what some have called our experiment with civilization. Populations became less nomadic. New social environments with hierarchies were created through the establishment of farming communities and towns. Mechanisms for the control of land, food production, irrigation systems, and organization of labor evolved to ensure that these more complex systems could work. These new social orders did not have the collective, communal aspects of the hunter-gatherer societies. As Boserup points out, many hunter-gatherer societies stayed away from the new order to protect their communal form of existence. It has been suggested that hunter-gatherer societies knew of the technologies necessary for agriculture long before the development of agriculture.

In the towns and cities that emerged, the quality of the water supply and sanitation systems became factors influencing health. The increased population density created extreme vulnerability to any new infectious agent that swept into the population. In addition, wars were largely fought over controlling the sources of food and its distribution. In this period, as today, war often contributed to famine and epidemics. Although medicine emerged as a profession during this period, the causes of disease were poorly understood and therapies were limited in their effectiveness. Despite these factors, there was steady population growth: from approximately 10 million in 8000 B.C. to approximately 750 million in A.D. 1750.

Towards the end of this period insight into socioeconomic policies and the health of populations was gained. In the sixteenth century the authorities in England recognized that famines were man-made rather than natural disasters. Food was available, but

the lower classes could not gain access to it.⁸ The state interfered and surplus food was supplied to the poor. This intervention angered farmers, producers, and merchants and contributed to the grievance against Charles I and the Civil War. After the revolution the new government abandoned the Tudor-Stuart program of food relief. This policy change subjected England to nearly two centuries of periodic famines. For the poorer members of the population, food riots became common. The issue was not food production, but distribution. After the Industrial Revolution, the government reinstated the old Tudor-Stuart program.

One of the criticisms that has been raised about the role of better nutrition in the improved health of populations is the observation that prior to the Industrial Revolution the peerage in England had the same mortality rate as the general population even though they had access to abundant supplies of food.⁹ One explanation is that an abundance of food did not ensure that the upper-class children had a better diet than the general population.¹⁰ After the Industrial Revolution, and its associated cultural and social changes, the life expectancy and height (a measure of the adequacy of nutrition during childhood) of the peers increased rapidly, while those of the general population lagged behind. Interestingly, a health gradient between the peers and the rest of the population emerged only after the Industrial Revolution.

THE INDUSTRIAL REVOLUTION

Although there were many technological innovations that affected socioeconomic development following the Agricultural Revolution, a major change occurred when it became possible to harness fossil fuels as a source of energy. Prior to this, slaves or serfs were used for most physical labor. Unfortunately, there is little data detailing the differences in the health and well-being of serfs and slaves versus those who were more privileged before the Industrial Revolution. As Rosenberg and Birdzell have emphasized, the Industrial Revolution was associated with vastly enhanced prosperity for Western societies, which led to better health, the disappearance of slavery and serfdom, and the development of democracy and universal suffrage.¹¹ During this period we find a substantial improvement in health and a dramatic rise in population. The world population has

increased from approximately 750 million in 1750 to nearly 6 billion today. The decline in mortality rates over the past 150 years is one of the great triumphs in human history. A UN report in 1953 attributed the trend to four factors: 1) public health measures; 2) advances in medical knowledge and therapeutics; 3) improved personal hygiene; and 4) improved standards of living.¹²

McKeown and Brown, attempting to better understand why health improved, explored factors influencing the major decline in mortality rates in the United Kingdom after 1840.¹³ They found that the bulk of the change in mortality rates could not be explained through medical intervention because there were no effective treatments for the major causes of death during most of this period. Furthermore, a large decline in mortality from airborne diseases could not be easily explained by the improved water and sanitation systems at the end of the 1800s. McKeown's conclusion was that the decrease in mortality was due to improved prosperity and nutrition.¹⁴

Critics pointed out that while McKeown's argument about improved nutrition was obviously part of the story, he gave too little credit to public health measures and medicine. He also did not account for discrepancies, such as why the mortality rate of infants and young children in the United Kingdom did not begin to fall until 1900 when the decline in death from tuberculosis started much earlier.¹⁵ Reves showed that the increase in child spacing that occurred in the United Kingdom around the beginning of this century was sufficient to increase the median age of exposure, and thus decrease susceptibility, to infectious diseases.¹⁶ In earlier societies, child spacing was, in part, driven by culture, population density, access to resources, and survival of the society. The dynamics of the practice of child spacing cannot be determined by a simple formula. However, increased birth spacing has been shown to have a positive effect on a nation's health.

Improved water systems and sanitation greatly decreased the number of deaths due to waterborne diseases. The greatest decline occurred during the latter part of the nineteenth century, when water and sanitation systems in Britain were improved. Also in the nineteenth century, the hygiene movement and its public health accomplishments¹⁷ reversed the deterioration of the British environment that was associated with the growth of urban centers during

the Industrial Revolution. The socially conscious citizens of an increasingly prosperous society drove their economic and political institutions to improve the urban environment, thereby improving health and well-being for the society.

Following McKeown's pioneering work, Fogel¹⁸ and others¹⁹ have provided new evidence about the relation between improved nutrition and changes in mortality rates during the Industrial Revolution. They shifted attention from famine (only a small factor in preindustrial mortality rates) to chronically poor or inadequate nutrition as a determinant of health. Before the middle of the nineteenth century, national food production in countries like England and France was not sufficient to provide adequate nourishment to the lower class. The bottom 20 percent of the labor force took in enough calories to stay alive but not enough to do much work.²⁰ In addition, undernourishment led to weakened immune systems and increased vulnerability to infectious diseases.²¹

Height and weight measurements have long been used as a means of better understanding the relationship between nutritional status and mortality. An individual's height in adulthood reflects the effects of the nutritional experience during the growing years, including the fetal period. Height at maturity is inversely associated with risk of chronic diseases and dying in the later stages of adult life.²² Weight represents the balance between nutrition in adult life and energy demands. The association of height with adult mortality rates reveals not the effect of adult nutrition but of nutritional levels (and disease history) from conception to maturity. Individuals who are poorly nourished or overnourished in adult life show higher mortality rates than individuals who maintain an ideal weight for height. However, short men who maintain an ideal weight are at a greater risk of death than are taller men.²³ Height is determined by genetic factors and nutrition, and the relative importance of each in explaining individual variation in height is still being debated. However, population mean height over time, which is used in these population studies, is primarily determined by environmental factors.²⁴

In Western countries, where records are available, the improvement in nutritional status during infancy and childhood, as estimated by changes in the mean height of the populations, is associated with a decline in mortality rates. Examination of English and American

data has shown that Americans achieved mean heights and levels of life expectancy by the middle of the eighteenth century which were not achieved by the British upper classes until the beginning of the twentieth century.²⁵ The evidence from these studies also shows that the lower classes in England did not show a marked increase in height until this century. Countries in which access to nutritious food varies by social class, for whatever reasons, tend to show class gradients in height and health status. In countries that have a high degree of income equity and equitable access to good food, height differentials by socioeconomic class have largely disappeared.²⁶

The historical records since the Industrial Revolution show fluctuations in mean heights in populations in Western countries that appear to be related to socioeconomic factors such as the state of the economy, income distribution, and the effects of urbanization. These studies of the relationship between poor nutrition during early life and health and well-being in later life have shown that chronic health problems are more common among short or stunted men than among tall men.²⁷ Rejection rates for recruits into the Union army in the US Civil War, based on medical conditions, were, coincidentally, inversely correlated with the height of the potential recruits.

Fogel²⁸ notes that individuals in the last century who survived into the later stage of adult life were not freer of chronic disease than are persons of the same age today. He makes the point that reliance on cause-of-death information has led to a significant misrepresentation of the distribution of health conditions and an unfortunate separation of the epidemiology of chronic diseases from contagious diseases. Poor development during early life, as a result of poor nutrition, not only increases the risk of dying from contagious diseases, it also increases the likelihood of chronic disorders in adult life. The evidence also indicates that this phenomenon is not disease-specific, but is related to the development of the immune system and other organ systems.²⁹

A historical analysis of the influences of economic growth and improved prosperity on health and well-being is constrained by the limited information available. Thus, while the case for improved nutrition, particularly during childhood, seems clear, the role of other factors, such as better nurturing of children, the environments in which individuals live and work, and other socioeconomic fac-

tors, cannot be readily determined. More recent studies have begun to show the roles of these other factors.

COMPETENCE, COPING SKILLS, AND HEALTH AND WELL-BEING

The strong and pervasive relationship between an individual's place in the social structure of society and his health status is striking. Kitagawa and Hauser showed compelling evidence of different rates of mortality according to socioeconomic class in the United States between 1930 and 1960.³⁰ For several major causes of death, the rates were highest for the lower social classes. Even though the mortality rates in the United States have continued to decline, the social gradient in health, as measured by levels of income and education, is still present and the differences in mortality rates have widened.³¹ This suggests that the widening in mortality rates is related to changing socioeconomic circumstances, such as increasing inequalities in income, education, and housing, a falling standard of living for a large segment of the US population, and limited access to health care for the poor and least educated.

Although life expectancy has improved for all social classes in the United Kingdom for the last sixty years, the gradient in social class mortality has been widening.³² In contrast, in Scandinavian countries the gradient in health has not widened and life expectancy has increased for all social classes.³³ It has recently been reported that the mortality rate for the lowest social class in Sweden is less than that for the top social class in the United Kingdom.³⁴

One of the best studies of the relation between socioeconomic factors and the health of the middle class is the Whitehall civil service study. This longitudinal study provides direct measures of the health of individuals against their position in a well-defined job hierarchy.³⁵ A striking finding from this study is the clear social gradient in health. As in larger population studies,³⁶ among those lower in the hierarchy there was found to be a higher risk of death from coronary heart disease, strokes, cancer, gastrointestinal disease, accidents, and suicides. The Whitehall study reported that the risk of dying of a heart attack for those in the bottom tier was more than 2.5 times that of the top tier. Marmot has shown that adjusting for conventional risk factors, such as cholesterol, blood pressure, and smoking, accounts for about 25 percent of the civil service

social gradient in coronary heart disease. The remainder of the risk is related to factors in the social environments in which those in the civil service live and work. What in the social environment influences our vulnerability to a wide range of diseases and has an effect equal to or greater than more conventional risk factors? Since the civil service is largely a middle-class population, the "something" that influences health is not affecting some underprivileged minority, but is affecting a larger population.

There are many beneficial medical interventions available today for sick or injured individuals. There are also a number of interventions of questionable benefit. In the Whitehall study, it was concluded that differences in medical care could not account for the three-fold differences in mortality among civil servants.³⁷ The Whitehall study also shows that life-style (e.g., smoking) is strongly influenced by an individual's position in the social hierarchy. The study also reinforced a key conclusion from the historical analysis: the mean height of the civil servants, in each job classification, correlates with position in the job hierarchy, sickness-absence rate, and risk of dying.³⁸

An individual's sense of achievement, self-esteem, and control over his or her work and life appears to affect health and well-being. Studies in Sweden have shown that individuals in high demand jobs who see themselves as having little control over their work have a much higher incidence of coronary heart disease than do people in similar jobs who believe they have control.³⁹ Similarly, the Whitehall study found that a high proportion of people in the lower tiers of the civil service feel they have less control of their work than do individuals in the top tiers of the civil service.⁴⁰

How competence and coping skills relate to vulnerability to disease may be explained by improved understanding of the links between the brain and the endocrine pathways and the immune system.⁴¹ We now understand some of the biological pathways through which individuals, coping with the demands of the environment in which they live and work, can influence the host defense system and disease expression. One set of animal experiments found that a friendly, supportive environment influences the process of diet-induced atherosclerosis. In these studies, two groups of rabbits were fed a cholesterol-rich diet. Those that were treated kindly (i.e., had music played to them) had 60 percent fewer incidences of

atherosclerosis than those given the usual laboratory treatment.⁴² Another research group working with monkeys found that an unstable social environment can accelerate coronary artery atherosclerosis, and that animals in the same colony, fed the same high cholesterol diets, show very different degrees of coronary artery occlusion depending upon their position in the hierarchy of the group.⁴³

Emerging evidence from fields such as psychology and the neurosciences points to how nurturing or stimulation influence brain development, particularly when it is most plastic.⁴⁴ The modifications and connections that are formed among the billions of cells in the cerebral cortex occur very rapidly during the first few months of life and continue throughout childhood. The development of the brain is strongly influenced by the quality of the nourishment and nurturance given to infants and children. The stimuli affect not only the number of brain cells in the cortex and the number of connections among them, but also the way the nerve cells are "wired." The stages in the development of the brain appear to be linked so that events early in life affect the development and function of the brain at later stages. In addition, in adverse environments activated stress hormones can have a negative effect on brain development and can damage neurons, leading to permanent defects in memory and learning.⁴⁵ Studies have suggested that children who were better nurtured in early life are healthier and do better in adult life.⁴⁶ There is evidence from studies in animals ranging from rats to nonhuman primates that show that this relationship exists.⁴⁷

This new understanding of the mind-body influence on disease expression also has relevance to some of the earlier observations. Could the post-1900 improvement in infant mortality in Britain⁴⁸ have been due, in part, to the link between breast-feeding and immune system responses, including the mother's early response to antigens in the infant's saliva? A mother, who is in circumstances that she has difficulty coping with and whose mind-body dynamics are suppressing her immune capability, may not be adequately building the defenses of the child she is feeding.

McKeown observed the steady decline throughout the nineteenth and twentieth centuries in mortality from tuberculosis, a disease which affects primarily children and young adults.⁴⁹ We know that tuberculosis is influenced not only by family age, structure, and crowding but also by host response capability, which may well be

affected by the social environment as well as adequacy of nutrition. Thus, it may be that the increased prosperity and control people had over their lives after the Industrial Revolution, accompanied by improved nutrition, increased the population's host defense capabilities and this inhibited expression of the disease. The relationship among the nervous system, the endocrine system, and the immune system is emerging as a pathway that can help our understanding of the changes in health which are associated with changing social and economic conditions.

People's positions in the hierarchy of a society, the degree of control they enjoy, and the adequacy of their diet appear to be important factors in determining vulnerability to a wide range of diseases. The relationship between the quality of nurturing and adequacy of nutrition in early childhood and health risks in adult life has implications not only for health policies but for policies concerning the competence and coping skills of the population (human capital). This is a key factor in determining economic growth. We need to better understand how economic forces influence the quality of social environments and human development. Countries with major improvements in health status and less inequality in health tend to be countries that are prosperous and have a high degree of social and income equity, and a small proportion of children living in poor social environments. Are societies that are more coherent or communitarian in character likely to provide better environments for health than are extremely individualistic, fragmented societies? How does economic growth and prosperity determine the quality of social environments?

ECONOMIC GROWTH, PROSPERITY, AND HEALTH AND WELL-BEING

In a recent essay on economic growth, *The Economist* said: "true enough, economists are interested in economic growth. The trouble is that, even by their standards, they have been terribly ignorant about it. The depth of that ignorance has long been their best kept secret."⁵⁰ A key issue has been the inability of the theoretical framework of neoclassical economics to cope with the role of technological innovation as an endogenous force in economic growth. The new framework of understanding of the determinants of economic growth, that embraces the role of technological innovation,⁵¹ makes

it important to understand better the relationships between technological innovation, economic growth, and the effects of these changes on society.

In his analysis of the improvements in the health of populations since the Industrial Revolution, Fogel points out that technological innovation can be a mixed blessing for populations that have to live through the associated socioeconomic changes.⁵² For example, there have been periods of technological change and vigorous economic growth which produced limited, if any, improvements in the health status of the populations.⁵³ As Sen has pointed out, how societies create and distribute their wealth determines the health and well-being of the population.⁵⁴ In periods of profound technological change, with associated changes in wealth creation and distribution, individuals, particularly the young, will be at risk.

Changes since World War II in Eastern Europe and Japan illustrate this relationship. The decline in the economies of Eastern Europe has been associated with a decline in the health status of the populations while the improved prosperity of Japan has been associated with a marked improvement in health status. In a recent analysis for the World Bank on the decline in health status in Eastern Europe, Hertzman concluded that, in addition to the deterioration in the physical environment, a strong factor seems to be the deterioration in the quality of the social environment in which families live and work.⁵⁵ In contrast, the extraordinary improvement in the health of the Japanese⁵⁶ is associated with enhanced prosperity and what appears to have been a remarkable ability to sustain the quality of social environments and reasonable income equity throughout the society.

Wilkinson has found in his analysis of a number of Western countries that life expectancy is correlated with the degree of income equity in the society.⁵⁷ A recent study of Northern England showed a widening inequality in health that was linked to increasing income inequality.⁵⁸ How societies create and distribute the resources necessary to sustain their populations is a fundamental question. Adam Smith concluded that there were sectors of the economy that produced the wealth that made other activities in society possible. Smith described these sectors in his chapter entitled "Of the Accumulation of Capital, or of Productive and Un-Productive Labour:"

The labour of some of the most respectable orders in the society is, like that of menial servants, unproductive of any value, and does not fix or realise itself in any permanent subject, or vendible commodity, which endures after that labour is past, and for which an equal quantity of labour could afterwards be procured. The sovereign, for example, with all the officers both of justice and war who serve under him, the whole army and navy, are unproductive labourers. They are the servants of the public, and are maintained by a part of the annual produce of the industry of other people. Their service, how honourable, how useful, or how necessary soever, produces nothing for which an equal quantity of service can afterwards be procured. The protection, security and defence of the commonwealth, the effect of their labour this year will not purchase its protection, security, and defence for the year to come. In the same class must be ranked, some both of the gravest and most important, and some of the most frivolous professions: churchmen, lawyers, physicians, men of letters of all kinds, players, buffoons, musicians, opera-singers, opera-dancers etc.⁵⁹

The new understanding of the factors determining economic growth has implications for Smith's splitting of the economy into productive and nonproductive labor. Since both sectors are important to a society, the productive labor section can be considered the primary wealth-creating sector (the engine of economic growth) and the other sector, the secondary wealth-creating component (the quality of the environment in which we live and work). When the primary wealth-creating sector falters, the income that flows to the secondary sector decreases, with associated changes in our social environment that can reduce our quality of life. The old economic theory tended to treat all outputs in the economy as being equal in wealth creation. The new concept clearly brings out the importance of a better understanding of a healthy primary wealth-creating sector and the synergy between this sector and the secondary wealth-creating sector. Many activities in the secondary sector, like some aspects of education, health care, and the support of children, are key parts of the infrastructure for all innovative economies.

Britain is regarded by many as a nation that has failed to make investments during this century in new technologies on the scale necessary to maintain its wealth-creating capacity, and its economy has fallen behind those of other developed countries.⁶⁰ An interesting question is how many of the inequalities in health in the United

Kingdom, particularly in the regions of major economic decline,⁶¹ are products of the failure to invest in the key technological innovations that determine economic growth?

Krugman makes the point that the United States has two major problems: slow growth in productivity and rising poverty.⁶² To increase the nation's wealth through enhanced productivity, ideas, and innovation is key. To help control expenditures, he points out, it is necessary to make the nation's health-care system more cost-effective. In a sense, we are back to the debate in the "Black Report" concerning the need to allocate resources wisely in the secondary wealth-creating sector of our economies. In Canada this debate has become part of the policy considerations concerned with health in the provincial governments.⁶³ In Manitoba, the new Centre for Health Policy and Evaluation has shown from its analysis of the records of their health-care system⁶⁴ (a proxy for the health status of the population) that there is a clear gradient in health against measures of social deprivation. It is not lack of access to health care that is setting this gradient but the underlying social economic factors (unemployment, income, and education) as in Marmot's study of the Whitehall civil service. As might be expected, the most deprived part of the population places the greatest demands on the health-care system. Manitoba is trying to confront the need for reallocation of resources, from the least appropriate health-care expenditures to the social needs of the population in greatest difficulty, particularly children in poverty. As Sen has pointed out, it is not the level of wealth a country has that improves the health of the population, but its commitment to allocate resources to key sectors, such as mothers and children, education, and adequate nutrition.⁶⁵

Canada, for the past twenty years, has not been creating sufficient wealth to sustain its consumption in the public and private sectors. Consequently, it has borrowed to maintain its standard of living, leading to the largest external debt per capita (private and public sector) in the developed world.⁶⁶ Canada faces the challenge of trying to sustain its social systems, including health care, social support, and education with diminished resources while simultaneously trying to rebuild the economy.

To sustain quality social environments with diminished resources is a difficult task. It is possible that societies with high-quality social

capital will be better able to adjust than will fragmented individualistic societies. Societies that have a strong, coherent sense of what is important, and a collective will, will probably be most successful. Putnam's description of what constitutes "civic societies" appears to be important in this regard.⁶⁷

It is time to integrate our understanding of the determinants of health and the determinants of economic growth. Governments and their societies are mistaken to concentrate on the economics of business cycles rather than the long-term forces affecting economic growth, prosperity, and health and well-being. Fogel has concluded that 50 percent of the economic growth in Britain since the Industrial Revolution has been due to better nutrition of the population.⁶⁸ A society that handicaps large segments of its population during periods of major technological change may be handicapping its future economic growth. We now have a better understanding of the relationships among economic growth, prosperity, and health and well-being, and the need for a long-term, integrated perspective on the determinants of health and economic growth. Can we make intelligent and wise use of this understanding?

ENDNOTES

¹William Beveridge, *Social Insurance and Allied Services* (New York: Macmillan, 1942).

²Great Britain Royal Commission on the National Health Service, "Report of the Royal Commission on the National Health Service (The Merrison Report)" (London: HMSO, 1979).

³Douglas J. Black, Cyril Smith, and Peter Townsend, *Inequalities in Health: The Black Report* (New York: Penguin Books, 1982). Since the contribution of socioeconomic factors was not a thesis the Thatcher government wished to have widely publicized, the government only allowed 260 copies of the report to be published. It was subsequently published by Penguin Books under the title *Inequalities in Health*.

⁴Ibid.

⁵Margaret Whitehead, *The Health Divide* (London: Health Education Council, 1987).

⁶Ester Boserup, *Population and Technological Change* (Chicago, Ill.: The University of Chicago Press, 1981).

- ⁷Lisa Berkman and S. Leonard Syme, "Social Networks, Host Resistance, and Mortality: A Nine-Year Follow-up Study of Alameda County Residents," *American Journal of Epidemiology* 109 (2) (1979): 186–204.
- ⁸Robert W. Fogel, "The Conquest of High Mortality and Hunger in Europe and America: Timing and Mechanisms," in David Landes, Patrice Higonnet, and Henry Rosovsky, eds., *Favorites of Fortune: Technology, Growth and Economic Development Since the Industrial Revolution* (Cambridge, Mass.: Harvard University Press, 1991).
- ⁹Robert W. Fogel, "Nutrition and the Decline in Mortality since 1700: Some Preliminary Findings in Long-Term Factors in American Economic Growth," in Stanley L. Engerman and Robert E. Gallman, eds., *Conference on Research in Income and Wealth*, vol. 41 (Chicago, Ill.: University of Chicago Press, 1986).
- ¹⁰The quality of early childhood nutrition can set health risks in adult life. See David J. P. Barker, *Fetal and Infant Origins of Adult Diseases* (London: BMJ, 1992).
- ¹¹Nathan Rosenberg and L. E. Birdzell, *How the West Grew Rich* (New York: Basic Books, Inc., 1986).
- ¹²The United Nations, *The Determinants and Consequences of Population Trends*, *Population Studies* no. 17 (New York: United Nations Publication, 1953).
- ¹³Thomas McKeown and R. G. Brown, "Medical Evidence Related to English Population Changes in the Eighteenth Century," *Population Studies* 9 (1955): 119–41.
- ¹⁴Thomas McKeown, *The Modern Rise of Population* (New York: Academic Press, 1976).
- ¹⁵Randall Reves, "Declining Fertility in England and Wales as a Major Cause of the Twentieth Century Decline in Mortality," *American Journal of Epidemiology* 122 (1985): 112–26; Simon Szreter, "The Importance of Social Intervention in Britain's Mortality Decline c. 1850–1914: A Re-interpretation of the Role of Public Health," *The Society for the Social History of Medicine* 1 (1988): 1–17.
- ¹⁶Reves, "Declining Fertility in England and Wales as a Major Cause of the Twentieth Century Decline in Mortality."
- ¹⁷Ibid. Anthony S. Wohl, *Endangered Lives: Public Health in Victorian Britain*, report on Public Health and Social Conditions (London: J. M. Dent, 1983).
- ¹⁸Robert W. Fogel, "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes in the Making of Economic Policy," working paper no. 4638 (Cambridge, Mass.: National Bureau of Economic Research, April 1994).
- ¹⁹Edward A. Wrigley and R. S. Schofield, *The Population History of England, 154–1871: A Reconstruction* (Cambridge, Mass.: Harvard University Press, 1981). Patrick Galloway, "Differentials in Demographic Responses to Annual Price Variations in Pre-Revolutionary France: A Comparison of Rich and Poor Areas in Rouen, 1681–1787," *European Journal of Populations* 2 (1986): 269–305. Zvi Eckstein, Paul T. Schultz, and Kenneth I. Wolpin, "Short-Run Fluctuations in Fertility and Mortality in Pre-Industrial Sweden," *European Economic Review* 26 (1985): 297–317. Alfred Perrenoud, "The Mortality Decline in a Long-Term Perspective," in Tommy Bengtsson, Gunnar Fridlitzius, and Rolf

Ohlsson, eds., *Pre-Industrial Population Change* (Stockholm: Almquist and Wiksell, 1984), 41–69.

²⁰Wohl, *Endangered Lives. Public Health in Victorian Britain*. Herman Freudenberger and Gaylord Cummins, "Health, Work, and Leisure Before the Industrial Revolution," *Explorations in Economic History* 13 (1976): 1–12.

²¹Ranjit Kumar Chandra, "Nutrition and Immunity: Lessons from the Past and New Insights into the Future," (1990 McCollum Award Lecture) *American Journal of Clinical Nutrition* 53 (1991): 1087–102.

²²Wohl, *Endangered Lives. Public Health in Victorian Britain*. Michael G. Marmot, Martin J. Shipley, and Geoffrey Rose, "Inequalities in Death—Specific Explanations of a General Pattern?," *Lancet* I (1984): 1003–1006. Hans T. Waaler, "Height, Weight and Mortality: The Norwegian Experience," *Acta Medica Scandinavica* suppl. 679 (1984): 1–51. A. Meridith John, *The Plantation Slaves of Trinidad, 1783–1816: A Mathematical and Demographic Inquiry* (Cambridge: Cambridge University Press, 1988).

²³Wohl, *Endangered Lives. Public Health in Victorian Britain*. Marmot et al., "Inequalities in Death—Specific Explanations of a General Pattern?."

²⁴Wohl, *Endangered Lives. Public Health in Victorian Britain*.

²⁵Fogel, "The Conquest of High Mortality and Hunger in Europe and America: Timing and Mechanisms."

²⁶Gro H. Bruntland, Knut Liestøl, and Lars Walløe, "Height, Weight, and Menarcheal Age of Oslo Schoolchildren During the Last 60 Years," *Annals of Human Biology* 7 (1980): 307–22.

²⁷Wohl, *Endangered Lives. Public Health in Victorian Britain*. Chandra, "Nutrition and Immunity: Lessons from the Past and New Insights into the Future."

²⁸Fogel, "The Conquest of High Mortality and Hunger in Europe and America: Timing and Mechanisms."

²⁹Barker, *Fetal and Infant Origins of Adult Diseases*.

³⁰Evelyn Mae Kitagawa and Philip M. Hauser, *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology* (Cambridge, Mass.: Harvard University Press, 1973).

³¹Gregory Pappas, Susan Queen, Wilbur Hadden, and Gail Fisher, "The Increasing Disparity in Mortality Between Socioeconomic Groups in the United States, 1960 and 1986," *New England Journal of Medicine* 329 (1993): 103–109.

³²George Davey-Smith, Mel Bartley, and David Blane, "The Black Report on Socioeconomic Inequalities in Health 10 Years On," *BMJ* 301 (1990): 373–77.

³³Whitehead, *The Health Divide*. Eero Lahelma and Tapani Valkonen, "Health and Social Inequalities in Finland and Elsewhere," *Social Science and Medicine* 31 (1990): 257–65; Jonathan S. Feinstein, "The Relationship between Socioeconomic Status and Health: A Review of the Literature," *The Milbank Quarterly* 71 (1993): 279; Denny Vagero and Olle Lundberg, "Health Inequalities in Britain and Sweden," *Lancet* II (1989): 35.

- ³⁴Feinstein, "The Relationship between Socioeconomic Status and Health: A Review of the Literature."
- ³⁵Chandra, "Nutrition and Immunity: Lessons from the Past and New Insights into the Future." Michael G. Marmot, "Social Inequalities in Mortality: The Social Environment in Class and Health," in Richard G. Wilkinson, ed., *Class and Health* (London: Tavistock Publications, 1986), 21-34. Michael G. Marmot, George Davey-Smith, Stephen Stansfeld et al., "Health Inequalities Among British Civil Servants: The Whitehall II Study," *Lancet* I (1991): 1387-393. Fiona North, S. Leonard Syme, Amanda Feeney, Jennifer Head, Martin J. Shipley, and Michael G. Marmot, "Explaining Socio-Economic Differences in Sickness Absence: The Whitehall II Study," *BMJ* (1993).
- ³⁶Vera Carstairs and Russell Morris, *Deprivation and Health In Scotland* (Scotland: Aberdeen University Press, 1991).
- ³⁷Chandra, "Nutrition and Immunity: Lessons from the Past and New Insights into the Future."
- ³⁸Ibid.
- ³⁹Robert Karasek and Tores Theorell, *Health Work: Stress, Productivity, and the Reconstruction of Working Life* (New York: Basic Books, 1990).
- ⁴⁰Michael G. Marmot and Tores Theorell, "Social Class and Cardiovascular Disease: The Contribution of Work," *International Journal of Health Services* 18 (1988): 659-74.
- ⁴¹Seymour Reichlin, "Neuroendocrine-Immune Interactions," *New England Journal of Medicine* 329 (1993): 1246. Robert M. Sapolsky, *Stress, The Aging Brain, and the Mechanisms of Neuron Death* (Cambridge, Mass.: The MIT Press, 1992).
- ⁴²Robert M. Nerem, Murina J. Levesque, and J. Fredrick Cornhill, "Social Environment as a Factor in Diet-Induced Atherosclerosis," *Science* 208 (1980): 1475.
- ⁴³Jay R. Kaplan, Stephen B. Manuck, Thomas B. Clarkson, and Robert W. Prichard, "Animal Models of Behavioral Influences on Atherogenesis," *Advanced Behavioral Medicine* 1 (1985): 115.
- ⁴⁴Carnegie Corporation of New York, "Starting Points, Meeting the Needs of Your Youngest Children," report of the Carnegie Task Force on Meeting the Needs of Young Children (New York: Carnegie Corporation of New York, 1994). Gerald D. Fischbach, "Mind and Brain," *Scientific American* 267 (1992): 48-57.
- ⁴⁵Reichlin, "Neuroendocrine-Immune Interactions." Kaplan et al., "Animal Models of Behavioral Influences on Atherogenesis."
- ⁴⁶Kristina Orth-Gomer and Jeffrey V. Johnson, "Social Network Interaction and Mortality: A Six Year Follow-Up Study of a Random Sample of the Swedish Population," *Journal of Chronic Diseases* 40 (10) (1987): 949-57.
- ⁴⁷Ibid. Stephen J. Soumi, "Adolescent Depression and Depressive Symptoms: Insights from Longitudinal Studies with Rhesus Monkeys," *Journal of Youth and Adolescence* 20 (1991): 273-87.
- ⁴⁸McKeown, *The Modern Rise of Population*.

- ⁴⁹McKeown and Brown, "Medical Evidence Related to English Population Changes in the Eighteenth Century."
- ⁵⁰"Economic Growth, Explaining the Mystery," *The Economist* (4 January 1992): 15-18.
- ⁵¹Richard G. Lipsey, *Globalisation, Technological Change and Economic Growth* (Ireland: Northern Ireland Economic Council Report, July 1993).
- ⁵²Fogel, "The Conquest of High Mortality and Hunger in Europe and America: Timing and Mechanisms."
- ⁵³*Ibid.*
- ⁵⁴Amartya Sen, "The Economics of Life and Death," *Scientific American* (May 1993): 40-47.
- ⁵⁵Clyde Hertzman and Wendy Ayers, "Environment and Health in Central and Eastern Europe," Report for the World Bank, No. 12270-ECA, 1993.
- ⁵⁶Michael G. Marmot and George Davey-Smith, "Why are the Japanese Living Longer?" *BMJ* 299 (1989): 1547-51.
- ⁵⁷Richard G. Wilkinson, "Divided We Fall," *BMJ* 308 (1994): 1113; Richard G. Wilkinson, "National Mortality Rates: The Impact of Inequality?," *American Journal of Public Health* 82 (1992): 1082.
- ⁵⁸Peter Phillimore, Alastair Beattie, and Peter Townsend, "Widening Inequality of Health in Northern England, 1981-91," *BMJ* 308 (1994): 1125.
- ⁵⁹Adam Smith, *The Wealth of Nations* (New York: Everyman's Library, Alfred A. Knopf, Inc., 1910).
- ⁶⁰Bernard Elbaum and William Laxonick, eds., *The Decline of the British Economy* (Oxford: Oxford University Press, 1986). Ralf Dahrendorf, *On Britain* (London: British Broadcasting Corporation, 1982).
- ⁶¹Michael Eames, Yoav Ben-Shlomo, and Michael G. Marmot, "Social Deprivation and Premature Mortality: Regional Comparison Across England," *BMJ* 307 (1993): 1097.
- ⁶²Paul Krugman, *Peddling Prosperity, Economic Sense and Nonsense in the Age of Diminished Expectations* (New York: W. W. Norton & Company, Inc., 1994).
- ⁶³"Nurturing Health: A Framework on the Determinants of Health" (Toronto: Premier's Council on Health Strategy, Government of Ontario, 1991).
- ⁶⁴Norman Frohlich and Cameron Mustard, *Socio-Economic Characteristics* (Winnipeg: Manitoba Centre for Health Policy and Evaluation, 1994).
- ⁶⁵Sen, "The Economics of Life and Death."
- ⁶⁶Peter J. Nicholson, "Competitiveness and the Canadian Economy" (Toronto: The Canadian Institute for Advanced Research, 1992).
- ⁶⁷Robert D. Putnam, *Making Democracy Work, Civic Traditions in Modern Italy* (Princeton, N.J.: Princeton University Press, 1993).
- ⁶⁸Fogel, "The Conquest of High Mortality and Hunger in Europe and America: Timing and Mechanisms."

A satisfactory national system for delivering primary or general medical care must have certain basic capabilities: (1) ready access to a physician or some other health professional who can effectively cope with ordinary medical problems; (2) the ability to identify, from innocent-appearing situations, those few that are potentially serious and provide properly for them; (3) science-based humanistic support for those that it serves; (4) care on a continuing basis; (5) distribution of care with reasonable equity; (6) the delivery system must be stable and self-renewing: those who work in it must find it sufficiently rewarding to want to stay in, and others must wish to enter, the field; (7) it must be compatible with the life, culture, and needs of the people it serves and adapt to the wide differences that exist in various areas of the country; (8) it must be able to compete effectively for resources with other social needs. Using these criteria, our primary health-care system is generally regarded as inadequate, and the supply of physicians specifically trained for, and willing to offer, primary health care as a significant part of their practice has been steadily declining for the last forty years.

David E. Rogers, M.D.

From "The Challenge of Primary Care"
Dædalus 106 (1) (Winter 1977)

Health Care as a Threat to Health: Defense, Opulence, and the Social Environment

WHAT "GOOD" IS HEALTH CARE?

VERY FEW PEOPLE OWN BATTLE TANKS for personal enjoyment. Such vehicles are not offered as prizes in lotteries or advertising programs or featured in the *Lifestyles of the Rich and Famous*. New homes, BMWs, tropical vacations, yes, but not tanks. And yet, under some circumstances, one might die if not for one then at least for lack of one.

The parallel with heart-lung transplants should be apparent. The services of the health-care system, like those of the military, are not wanted for their own sake. When they are called on, they are *needed*, and the need may be a matter of life or death. While in both cases having the services when they are needed is much better than not having them, it is far better still not to need them at all. The most sophisticated and effective health care in the world cannot produce results as good as simply remaining healthy in the first place. This is obvious to anyone who has ever been a patient—or been in battle—and our common understanding is pretty clear on the value of both health and peace.

Yet this understanding becomes clouded when we consider health-care policy because most health-care services (unlike those of the military) are personal services provided to individuals. By a progressive elision, a modification of the metaphor, the observation of a discrete commodity or a service provided to and utilized by an

Robert G. Evans is Fellow and Director of the Program in Population Health, The Canadian Institute for Advanced Research, and Professor, Department of Economics, The University of British Columbia.

individual can be interpreted as a "good" being "demanded" by a "consumer" rather than as an intervention being undergone by a patient. The overtones of uncertainty, anxiety, and suffering which accompany the experience of health care are replaced by the confidence and satisfaction of a knowledgeable (and affluent) buyer making choices in a well-stocked market.

In introducing this alternative set of terms, with its heavy load of (untested, rather implausible) behavioral assumptions and (personal, rather idiosyncratic) values, the naive economic analyst can dissolve the difference between "needing" a surgical procedure or a diagnostic intervention and "demanding" an ice-cream cone or a new pair of shoes. The same analytic apparatus is then applied to each type of transaction; one should not be surprised when similar policy recommendations emerge.

But, as *The Economist* said about theories of economic growth, it takes many years of training to ignore the obvious.¹ Most people do not have that training and may not have the intellectual capacity to absorb it in any case. Thus they have not lost their ability to perceive the obvious fact that health care is not a "good" in the same sense that a hamburger is a "good." When we hear that one person "consumed" an expensive vacation in the south seas and another person "consumed" an equally expensive episode of hospital care, we all know that the first event is to be celebrated (or envied) and the second is to be commiserated (or pitied), even if the person who took the vacation paid the full cost and the hospital patient received her large allocation of "goods" for free. A successful *outcome* of the second event is cause for celebration, but the episode in total is not.

COLLECTIVE FINANCING AND COLLECTIVE PRIORITIES

This broad public understanding is reflected in the fact that all countries in the developed world fund health care collectively, and almost all fund it through public or quasi-public agencies. *Table 1* shows the proportion of total health-care funding for the Organization for Economic Cooperation and Development (OECD) countries that comes through the public sector, defined broadly to include "social insurance."

TABLE 1. International Comparison of the Public Share of Health-Care Spending

	1980	1990
	(in percentages)	
Australia	63%	68%
Austria	69	67
Belgium	83	83
Canada	75	73
Denmark	85	83
Finland	79	81
France	79	74
Germany	75	73
Greece	82	76
Iceland	88	87
Ireland	82	75
Italy	81	76
Japan	71	72
Luxembourg	93	91
Netherlands	75	71
New Zealand	84	82
Norway	98	95
Portugal	72	62
Spain	80	78
Sweden	93	90
Switzerland	68	68
Turkey	27	36
United Kingdom	90	84
United States	42	42 (48)*

*adjusted to include tax subsidy

Source: George J. Schieber, Jean-Pierre Poullier, and Leslie M. Greenwald, "U.S. Health Expenditure Performance: An International Comparison and Data Update," *Health Care Financing Review* 13 (4) (Summer 1992).

Even the exception, the United States, the home of private finance, is less exceptional than it appears at first glance. Although nearly two-thirds of the *population* rely on private insurance for their primary coverage, this only covers about one-third of total *expenditures*, compared with 40 percent from public sources. And private coverage depends on heavy public subsidies through the income tax system. If these are netted out, "private" coverage accounts for only about one-quarter of the costs, compared with roughly half from the government.

The form of collective financing greatly affects how the total burden is distributed among the population. When health-care systems are funded from general tax revenue, people with higher incomes pay a significantly greater share of the overall costs than they do when health care is funded through specially earmarked "social insurance premiums."² Private insurance, along with user-pay or co-insurance charges, takes a particularly large proportion of the incomes of lower income people.³ Moreover, systems which rely more heavily on user payments and on multiple forms of private insurance also offer people with higher incomes better access to services—they pay less and get more. Regardless of how the burden is distributed, however, the predominance of collective funding is universal in the developed world.

The issue of who pays and how rarely arises in the financing of military services. But debates over health-care finance continue because health care is linked with specific persons and because it has been misleadingly labeled as a "good." The fact that the different ways of organizing the collective funding of health care have such very different distributional consequences, and that changes can have the effect of transferring substantial amounts of money from one set of pockets to another, assures that the conflict over funding will never be far from the top of the public agenda in most developed societies.

Moreover, it is not difficult to see that a very small proportion of the population—those unfortunate enough to be quite ill—account for most of health-care spending.⁴ For this reason, all systems rely on collective funding. In the modern world there is no possibility that those who have the greatest need for the services could ever pay a significant share of the cost from their own resources, even if we considered that fair.

Private insurance, operating on true insurance principles and without heavy public subsidy, is equally incapable of covering more than a small share of total outlays, though private insurers will happily enroll large numbers of people with minimal needs. The greatest flaw in the private insurance approach is that they are able to do so. Use is not only very unevenly distributed over the population, but it is also, to a large extent, predictable. At the same time, however, the fact that most people use very few services, and a few

use most, necessitates the large interpersonal transfers of resources that are an inevitable part of any modern health-care system.

By contrast, it is difficult to say exactly who in a society benefits most from military services. Awkward questions about who is paying for whom do not arise in so transparent a form. There is thus little controversy over funding from general tax revenue, the necessity of setting collective, "national" priorities to determine how much to spend in the pursuit of military objectives, or how to spend it. That is not to say that this collective decision-making process is not highly contentious or subject to extreme uncertainty. But it is generally accepted to be properly a collective, political debate without the intrusion of the individualistic rhetoric that so confuses public debates over health-care policy.

WHO ARE THE BENEFICIARIES? OUTPUTS AND INPUTS

In the military and in medicine, benefits (and costs) come in several different forms. There is a particular confusion between those who benefit from the services—people protected from invasion or cured of illness—and those who benefit from the opportunity to *provide* the services—doctors or career officers, defense contractors or drug company shareholders. Indeed, it is often argued that *everyone* gains from the expansion of these sectors through the creation of jobs. Particularly in times of high unemployment, one finds health care celebrated as a growth sector even while it is being criticized for absorbing an ever larger share of national resources.

The financial press is particularly prone to this confusion. They tend to interpret as "news" the competing claims made by the various parties who have a direct financial interest, as either payers or payees. Rarely is any effort made to assess the validity of those claims; they are simply juxtaposed with counterclaims.

There is no doubt that payees—not just doctors, nurses, and drug companies but also (among others) managers, health insurers, and trial lawyers—benefit handsomely from health spending. Every dollar of spending translates into a dollar of income, regardless of whether it buys any additional health or happiness. Useless or harmful services will generate income just as effectively as those that are beneficial—which is why quack medicine never dies out. Those for whom health expenditure represents income routinely

proclaim that what is good for them is good for America. And the media duly report their claims, more or less at face value.

Economists tend to have a clearer grasp of the distinction. No one has ever improved on Keynes' example that if the Treasury were to bury bottles full of bank notes (secretly, at night), a highly profitable private employment could be created as people went out to look for them and dig them up. Jobs would be created. Yet, no one would imagine that the average real output, income, or well-being of the United Kingdom would be increased. The gains of successful finders, their increased claim on national output, would be offset by the reduced share of those who did not find bottles. Indeed, if the bottle-finding industry absorbed not only the unemployed but some who would otherwise have been producing commodities of real value, then average incomes in real terms would in fact fall. It is the *output* that matters, that creates value or wealth. If there is no (useful) output, the job or activity is not of value in itself.⁵

The same idea is expressed in the more recently coined phrase "real jobs." The collateral suggestion that "real jobs" are found only in the private sector, and "unreal jobs" in the public, is ideological nonsense; the American health-care financing system, for example, currently wasting about \$100 billion a year in activity without (net) output, is probably the largest single concentration of "unreal jobs" in the world.⁶ But the distinction is real enough. To maintain living standards, a society must concentrate its scarce resources in activities that create wealth, not waste motion.

WEALTH THROUGH PRODUCTION, NOT DESTRUCTION

Again a military analogy may help clarify the point. Imagine two countries engaged in a major war that, to separate the human factors from the economic, is being fought only with and against "smart" weapons, not human beings. National income accountants on both sides dutifully record the production of various weapons platforms and ordnance in the Gross Domestic Product; these are then destroyed in battle the next week. Employment and output in each country are at an all-time high, as are recorded incomes. But is this prosperity? Keynes' bank notes in bottles have been updated to the ultimate video game.

The reality is that, on average, citizens in both countries are receiving a large part of their real “incomes” in the ephemeral form of military equipment. To reflect the rather limited value to them of this type of income, particularly in its postcombat form, their disposable monetary incomes must be reduced by a corresponding amount. This can be done through contemporary taxation, but typically wars are financed, in part at least, through borrowing. Inflation subsequently reduces the value of the loan, resulting in de facto taxation of the lenders.

One way or another, the diversion of resources into “products” that are not valued in themselves (and in any case do not last) must be matched by a reduction in the availability of commodities that people *do* value, either for consumption or as productive investments. Elementary economics texts often refer to a choice between “guns and butter”—more of one, less of the other. And the costs of military spending may not just be reductions in current consumption. If resources are diverted away from maintenance, repair, and upgrading of the nation’s capital stock (except for that needed for war production), long-term economic growth will suffer.

This is not to argue that the military activity itself is wrong-headed; it may be eminently sensible policy for one or even both countries. But the “income” and “output” generated in the process do not create wealth or prosperity. Furthermore, given the economic costs of war, people on both sides might ask whether the ultimate objectives of peace, security, and national fulfillment might possibly be pursued in a less costly manner by other means. Recall again that the production and operation of weapons have (for most of us) little or no value in themselves, but are means to these real, though poorly defined, ends.

ILL HEALTH IS NOT WEALTH

To bridge to the health-care system, consider a nation that is faced with a serious epidemic (for example, an HIV mutation that is transmissible through airborne droplets; Bovine Spongiform Encephalopathy [“mad cow disease”] that infects humans on a large scale; multi-drug resistant tuberculosis, etc.). If the problem were severe enough we would presumably see a massive transfer of resources into the health-care sector for treatment, prevention, and

research. One might predict that health-care cost containment would rapidly lose its position on most national agendas.

This major expansion of activity in the health-care sector would be recorded as national income and output. And, assuming that the resources were being deployed in a rational manner, given the available information most of us would probably accept that the national effort was “needed.” But, as in war, no one should be confused by the economic statistics or by the increased health-care activity into believing that we had become wealthier or healthier. Indisputably, such an epidemic would make us worse off, on both counts.

But the numbers can be very deceptive. Suppose the epidemic tended to carry off the very old and the very young—not an uncommon situation—and that treatment, despite heroic efforts, was not very effective. The net effect of the epidemic would be an *increase* in per capita output and income. The less effective the treatment and the heavier the death toll, the greater the increase.

Such an epidemic is, of course, purely hypothetical, although the scenario is less farfetched than it would have seemed twenty-five years ago. In developed societies the health-care sector devotes a very small share of its activity and resources to dealing with infectious diseases; even the extraordinary phenomenon of HIV/AIDS is highly localized in its impact. Attempts to recruit AIDS treatment as an explanatory factor in the escalation of health-care costs in the United States have no empirical support and are merely part of a broader effort to divert attention from the real forces at work.⁷

Nevertheless, the hypothetical epidemic highlights characteristics of the broader health sector. It illustrates very clearly the status of health care as *not* a good in itself but as a means, sometimes powerful but often imperfect, to an end, namely health. There are certainly circumstances in which health care, like military preparedness, is desperately needed by individuals or by the whole society. But those circumstances are defined by the existence of threats to health, against which care is effective; while we may *need* care, we do not *want* (to need) it.

COLLECTIVE COPING WITH THREATS—QUIS CUSTODIET...?

Health-care systems, like the military, embody our collective responses to perceptions of need. Threats to health and national

security are interpreted, and responses are formulated, by complex and highly professionalized establishments in each case. And the status, incomes, and professional self-fulfillment of the people in those establishments depend upon the level of collective support they receive from the rest of us. This in turn depends upon our perception of the seriousness of the threat and the effectiveness of the responses which they can offer.

It is no accident that the self-perpetuating "military-industrial complex" that President Eisenhower identified and warned against just before leaving office was so quickly transposed as the "medical-industrial complex"—the large, interdependent collection of professional and commercial interests, people and institutions, who offer protection against threats to health. In return, they require that we tithe—they take between 8 percent and 10 percent of the national incomes of most developed societies and over 14 percent in the United States.

Our ancestors tithed in hope of life after death; we hope for more and better life before death. Now, as then, those who collect the tithes offer hopes not promises. One must also live a virtuous life, and then perhaps. Nor do the professionals encourage or welcome close questioning about how the money is spent: for the greater glory of God, the advancement of science, the relief of human misery—the rest is not our business.

Of course, health, like peace and security, is well worth defending. But the same questions arise in each case: 1) How serious is the threat? 2) How effective are the responses offered? 3) How efficiently are those responses being mounted? 4) Are there other ways we could respond? 5) What are we giving up in the process? These questions do not bear on nonsense issues such as whether or not we, as a society, should support a health-care system or a defense establishment. Of course we should, and will. The real policy issues are how large they should be and what they should do. When such issues are left to the professionals, the answers are, invariably, "Somewhat bigger" and "Whatever *we* think appropriate." Implicitly the answers they offer to the five questions above are "Very," "Very," "Very," "No," and "Nothing of comparable value."

Well, the threats *are* real enough, and some of the responses are outstandingly effective. Few would be so foolish as to deny the powerful contribution that health care can make. But there are also

numerous examples of ineffective and inefficient responses, detailed, for example, among the various ways in which one can have “too much” medical care. These illustrate yet again the force of the ancient question, “*Quis custodiet custodies ipsos?*” Who keeps the professionals up to scratch? They cannot do it themselves.

Yet, during the quarter century following World War II, the peoples of the developed world supported the development and expansion of systems that were based on the implicit assumption that they could. If the professionals were given the tools, they would finish the job. They would identify the needs, define and apply the appropriate responses, and cure our ills—and that would be that. And indeed it was in some instances—for example, tuberculosis or polio.⁸

But it was soon discovered, in all countries, that if the professionals are asked to define “needs,” they will expand the definition without apparent limit. As one need is met, another will be found to take its place. The apparently reasonable and humane statement “we must meet the needs” becomes, in effect, a demand for an ever-expanding share of national wealth. In 1977 Aaron Wildavsky expressed this as the Law of Medical Money: “. . . costs will increase to the level of available funds” thus “. . . that level must be limited to keep costs down.”⁹

In response, all societies with modern health-care systems developed, at differing times during the 1970s, externally applied administrative mechanisms to contain the growth of their health-care systems within acceptable boundaries. The only exception, the United States, continues to confirm Wildavsky’s generalization. Health-care expenditures in the United States are well on their way to 20 percent of American national income.¹⁰

The Law of Medical Money applies to aggregate spending. Within the aggregate, however, the common experience has been that professionals do not, in fact, select the most effective and efficient ways of responding to the needs that they perceive and interpret. Rather, and perhaps unsurprisingly, they tend to choose responses that also advance their own interests in autonomy, professional self-expression, and income. Thus, systematic patterns of questionable or inappropriate care, inefficiently provided, have persisted for decades. Research findings to this effect have steadily accumulated in medical literature. But there are no *internal* pressures to respond.

Such patterns are, from a professional perspective, not a problem but a solution.

The administrative constraints on costs that have been relatively successful outside the United States have had a more mixed record in addressing these problems “under the aggregate.” The near-universal view at present seems to be that “better management” is the key to greater value for money. But the United States has been in the vanguard in demonstrating yet another lesson: clinicians are not the only professionals.

Managers have become another powerful set of claimants in the health-care sector, adding their own ambitions to those of the providers of care. The proportion of the resources devoted to health care in the United States that is absorbed as “administrative overhead”—in hospitals and clinics, payment agencies, and various forms of business services—has exploded in the last twenty years and continues to escalate rapidly.

If health care is not of value in itself, but only as a means to health, then administrative activity is twice removed from the things that matter. It has no value except insofar as it facilitates the provision of health care. Increases in overhead costs which cannot be shown to correspond to increases in efficiency or effectiveness represent waste, pure and simple. But to managers, who may, like Keynes’ bottle-hunters, work long hours with skill and dedication, they represent income.

Management is getting things done through other people.¹¹ But what things? If health-care managers are free to choose their own ends, as clinicians have been, then they will manage in their own interests, just as clinicians do, and perhaps with a good deal less concern for the patient. Those interests will not include global cost control and will have no more than an accidental correspondence with improved health outcomes.

So, management itself must be held accountable, but to whom and for what? In most of the world, management is accountable to a duly constituted and effective political authority, which pushes the accountability question back one stage. Aware of the inadequacies of *that* process, a number of Americans instead look to a (properly structured) competitive marketplace to force professionals to attend to the interests and concerns of “consumers.” But no one has ever seen this trick performed, and all efforts to date to

achieve accountability at a system-wide level through market mechanisms have failed. The faulty identification of health care with “goods” and patients with “consumers” underlies much of the misplaced optimism about what may be expected of markets in this sector.¹²

GRASPING SHADOWS, LOSING SUBSTANCE?

But how serious is the threat? What exactly *is* the threat? Death? Disability? Pain and suffering? In some absolute sense, death *cannot* be a meaningful threat since it is an ineluctable consequence of life. “Mortality is the price of identity.” Premature, agonizing, terrifying death—these are threats. But death just *is*. As Eliot’s Becket says: “I am not in danger; only near to death.” Similarly, functional disability is a side effect of aging—people who are eighty cannot do many of the things they did at twenty. The process of deterioration can be slowed, and some deficits repaired or compensated for—for a time. Some capacities improve with age, but again for a time. But perfect health, indefinitely maintained, like perfect security, is not on offer. “Here is no continuing city.”

If we misread the possible, we misperceive the threat. The result can be, often is, wholly inappropriate responses. During life, any discrepancy from the ideal must be met with some intervention, regardless of the chances of success. Raising and contributing the necessary resources is then presented as an ethical obligation on the rest of society. *Whose* ethics demand this? What are the grounds for obligation?

At some level we all understand the foolishness of this, just as we understand that increased defense spending, if it triggers an arms race, can easily *reduce* national security. Indeed we now have a further ironic twist. One interpretation offered for the highly-publicized American arms buildup of the early 1980s is that, whatever the underlying reality, it forced the Soviet Union to try to keep pace, thereby over-straining an already weak economy to the point that the regime collapsed. Clearly, building more tanks did not make the Soviet Union more secure; the world’s largest tank armies, never committed to battle, are now a major source of high-grade steel scrap.

But are the United States and its allies now more secure? That seems far from clear. Are *individual* Americans more secure now than they were ten or twenty years ago, either economically or physically? And what was the cost, in terms of lost opportunities?

THE DETERMINANTS OF HEALTH: MORE WAYS TO SKIN A CAT

The understanding that health, of both individuals and populations, is powerfully influenced by factors other than health care is ancient. It has been somewhat eclipsed, over the past century, by the extraordinary effect of the application of science and technology to medicine. But it is now enjoying a considerable resurgence, not only because of the development of a more measured and critical view of the contributions of health care but also because of the research findings on the determinants of health.¹³ Some of these are coming from sociology and psychology, always a bit skeptical of the claims of “mainstream” medicine. But surprisingly supportive evidence is also emerging from epidemiology, molecular biology, neurophysiology, and even primatology.

Possible biological foundations are beginning to emerge for the widely observed correlation between prosperity and health, a correlation which includes but goes far beyond the obvious connection between poverty and deprivation and corresponding physiological damage. There is extensive evidence for a correlation between health and hierarchical position within populations. This may reflect the effects on health of differences in material circumstances, education, status, or sense of control over one’s environment. If the relationships between health and “other things”—prosperity, self-esteem, sense of control—within populations also underlie the differences between populations (either different societies at the same time, or “the same” society at different times), then, quite conceivably, economic growth and development in itself is a powerful contributor to health.

But the connection is not simple. At a very low level of development, economic growth may mean more and better food, housing, clothing—(relatively) well-understood pathways to better health. But in more highly developed societies, the connections may be more subtle. Does prosperity per se contribute to health, or is it a marker for other forms of progress? Economic output is an indica-

tor of what a population can do. But it is not the only one. Ability is a powerful source of self-esteem. But it is not the only one. So the pathways are not clear. But it *is* clear that such pathways exist. The connections between health and social environment, broadly defined, are very real and may be very powerful.

OVERSHOOTING THE MARK

This raises the possibility that overinvestment in activities which are not themselves directly valued—such as medicine and the military—could have negative effects on health through the “opportunity costs” which they impose on the rest of the social environment.

The withdrawal of resources from more productive activities in order to build tanks and aircraft that then destroy, and are destroyed by, other tanks and aircraft¹⁴ is not difficult to perceive as a loss of national wealth. But suppose we (think we) are winning? Our sense of self-esteem and control may be enhanced even while our standard of living is being reduced. If psychological state is more important for health than is material wealth (at least above some minimum standard), then health may be improved even while prosperity is reduced. The perception of victory probably cannot be indefinitely sustained on both sides, however, and even the victors must count the costs on the morning after.

The transfer of resources to the health-care system is less clear-cut in its effects because health care may generate significant benefits both as consumption—improving the quality as well as the length of life—and as investment—restoring people to productive roles in society. But the efforts to contain the expansion of health-care systems, which are virtually universal throughout the developed world, are *prima facie* evidence that the members of those societies, or at least their political representatives, do not regard the benefits, at the margin, as worth the costs. In the United States, one hears the concern that excessive health-care costs are weakening the nation’s ability to compete in world markets, thus threatening the prosperity (and perhaps the health) of its population.

EASY ANSWERS DO NOT LIGHTEN THE LOAD

On a superficial level, economic analysis has a ready response. A nation *cannot*, in general, “price itself out of world markets.” If

domestic costs of production go up, for whatever reason—higher wages or higher health-care costs—a downward adjustment in the exchange rate can always restore the balance. And if American workers choose to take more of their compensation in the form of health-care benefits, and less as take-home pay, why should that be a source of public concern? If health care is a “good,” then “consumers” are simply substituting one “good” for another, presumably because that gives them greater satisfaction.

But this superficial response is weak at two points. If health care is *not* recognized as a “good” in the usual sense, then the transfer of an increasing share of national income to its purchase, particularly in the absence of any observable off-setting increases in health, will be quite correctly perceived by workers, and payers of insurance premiums and taxes, as a decrease in their real incomes. “Too much” health care is being produced at “too high” a price through processes over which the individual, the employer, and the insurer have, in practice, little or no control. Hence the distress.

The threat to international competitiveness may then arise as a consequence of people’s efforts to maintain their real standard of living, net of increasing health-care costs, whether those costs are channeled through taxes, payroll deductions, premiums, or direct charges. If these efforts result in lower rates of saving—expressed in lower personal savings, reduced corporate retained earnings, and larger government debt—this in turn reduces the resources available for investment in physical plant and research and development.

In wartime, patriotic propaganda (or fear) may encourage people to accept higher taxes and increase savings (buy Victory Bonds), while rationing can be used to reduce the availability of consumer goods. But no such mechanisms can be used to win acceptance of the transfer of resources from personal consumption to health care.

A country then becomes “uncompetitive” not because it *prices* itself out of world markets, but because its *products* are no longer competitive. No one wants products that embody obsolete technology. The second flaw in the economic analysis is that it implicitly assumes a fixed set of commodities and an unchanging technology. Only prices change, so any increase in domestic price levels can be canceled out by an adjustment in exchange rates. But the world we live in is dynamic. It is true—it has to be—that eventually exchange rate adjustment (or default) will bring a country’s international

accounts into balance. But in the process, its capacity to create wealth may have been severely damaged.

OLD IDEAS: NEW EVIDENCE AND APPLICATIONS

The argument that too much health care may be a threat to health, just as too much military effort may be a threat to national security, links together very old and very new ideas in both biology and economics. The ancient understanding that how we live has profound effects on our health is being substantiated and made much more specific—although also much more complex—by recent advances in both the biological and the social sciences. If an overdeveloped health-care system drains away resources that might otherwise have improved the social environment, it could be considered a threat to health.

At the same time, the most recent attempts to understand economic growth—the advancement of the wealth of nations—is reviving the concerns of Adam Smith and the classical economists who wrestled with the same problem. Their distinction between “productive” and “unproductive” labor (dismissed as wrong-headed by the marginalists over a century ago) points towards the modern concern with wealth-creating versus wealth-dissipating activities.

Quite obviously, activities which yield no output, or output that is of no value, are wealth-dissipating. But even valued outputs may be “unproductive” if they tend to reduce the capital base of a society. Modern theories of economic growth focus more on the importance of intangible capital—“know-how” or “science-based innovation”—than on physical capital like canals, railways, or factories. But know-how is commonly embodied in, and indeed arises from, investments in new physical capital.

Adam Smith would not have been surprised to find that a society that devotes a relatively high proportion of its economic resources to “unproductive” activities has a lower rate of growth. He was not a pacifist—he held that “defence is more important than opulence”—and he would presumably have included defense against nonhuman threats to life and limb. But the associated activities, however necessary, do not contribute to economic growth.

The linkage of military and biological defense suggests a more modern metaphor for the argument sketched out in this paper.

Each of us is defended against external threats by an immune system whose extraordinary complexity and sophistication is only beginning to be revealed. Without it, we could not survive a day. But that system can be stimulated to overreact to threats, or to react to nonexistent threats, in such a way as to become itself a larger threat to health. We are also protected against external threats at the collective or population level; the military and the health-care systems are important components of the social "immune system." But those systems too may malfunction.

ALLERGIES AND AUTOIMMUNITY IN SOCIAL IMMUNE SYSTEMS

Military metaphors have always come naturally to health care—"the battle against disease and death"—and particularly to descriptions of the immune system. Patrolling cells identify foreign invaders, report their type and location, and rally armies of compatriots to destroy the invaders along with native cells that have been subverted. Some cells attack directly, others use chemical weapons to neutralize at longer range, still others produce new fighters, and so on.

This protective activity is not limited to defense against foreign invasion. The immune system of the organism also responds, in ways that are significant though less well understood, to prevent or repair injuries and malfunctions. But immune responses also draw on a substantial share of the energy of the organism. Fighting off invaders or repairing domestic damage is not cheap, and even when successful it can have serious consequences for normal functioning. Infections can leave one not only prostrate during the acute episode, but weak and tired long afterward. The immune response competes particularly strongly with the normal processes of growth and development, maintenance and repair, and reproduction, which are also very expensive for the energy budget of the organism.

Survival requires that the response to immediate threats take precedence over longer term "capital investments." But if the response is prolonged, the development of the organism will suffer. Extended exposure to infection during critical stages of development may lead to permanent stunting even in children with apparently adequate diets because the available energy is diverted to defense, not growth.¹⁵ While the question of whether the deploy-

ment of the social “immune systems”—military and medical—can have negative effects on the growth of the economy is still open to debate, at the level of the organism it appears to have been answered.¹⁶

If the threat is real, then the price must be paid, whether in military or medical spending, or in energy used up by the immune system. But each system is highly complex and sophisticated, and each is subject to internal malfunction. In particular, each requires the capacity both to respond rapidly and effectively—which may mean massively—to external challenge yet also to shut down, quickly and fully, when it is no longer needed.

The immune system achieves this through a complex web of positive and negative feedback loops, which can rapidly expand and strengthen or dampen and suppress its responses. If these malfunction, however, the system may misinterpret the “self” and “nonself” signals and attack the organism’s own tissues. The list of resulting autoimmune diseases seems to be expanding along with knowledge of the immune system itself.

The immune system may also respond inappropriately to external, “nonself” stimuli—overreacting to pollen, or MSG, or bee venom, for example—with effects that are far worse than those of the triggering substance. Like autoimmune responses, allergic reactions can range from mildly troublesome to fatal. In both cases the immune system itself has become the threat.

Not dissimilarly, overdevelopment of the health-care system increases the possibility of iatrogenic illness, the illnesses *caused* by medical intervention. And too large a military can be a threat to peace and security both through the reactions it induces in other countries and through its increased presence in the home country. The military coup is an extreme form of autoimmune disease. These potential negative effects are additional to the inevitable impact of defensive systems, at the individual or societal level, in drawing resources away from other uses.

The major difference between individual and social immune systems is that in the former negative feedback loops have evolved such that the system normally turns itself off when the threat is past. As noted above, the social immune systems have no such automatic, internal controls, and must be restrained by deliberate external actions. These, however, are politically contentious—think

of the controversy that arises when the decision is made to close a military base or a hospital. Like Keynes' bottle-burying program, activities which create no wealth in aggregate nevertheless create considerable wealth for some. When the gainers are both concentrated and self-aware, and the losers are neither, the politics are not particularly difficult to understand.

But diffuse costs are no less real. They may show up in the present as a sacrifice of other sources of satisfaction, or in the future as reduced capacity to create wealth. This future capacity requires present investments in the form of physical capital (plant and equipment, roads and bridges), or human capital (education), or know-how (research and development). If these are sacrificed because the members of a society attempt (successfully) to maintain their present standard of living in the face of rising medical (or military) claims on total output, then in the long run growth of the economy will be stunted.

To one who views health care as a "good," such behavior would be irrational. From this perspective living standards are not reduced when less is spent on food, housing, or recreation, and more on medical care. One set of goods is being substituted for another, presumably because "consumers" want it that way. Why should they react by reducing savings, either personal or through government or corporate budgets?¹⁷ But from the alternative perspective offered here, it is entirely understandable and indeed predictable that people should fail to count increased health-care outlays as increases in their living standards and try to protect those standards *net* of such outlays.

There are at least some grounds to suspect that this may be happening in the United States. Savings rates are unusually low and health-care spending is unusually high and rising. And there is a major struggle among business, labor, government, and consumers as to who shall bear the brunt of cost increases—a struggle that diverts all participants from more cooperative efforts at control. Very clearly, American workers do not count increased health-care spending as part of their real incomes, whatever economists may tell them about the mix of pay and benefits in labor compensation.

The American health-care system represents an extreme form of the hyperreactive social immune system, responding ever more forcefully to a wider and wider range of perceived "threats," until it may

itself become a threat to health. But every society faces the problem of containing the steady pressure for expansion of health care. The American system has gone out of control because it lacks the structural features—universal coverage and political accountability—that have permitted some degree of containment in other countries. Nowhere, however, do we find the internal, self-stabilizing behavior of biological immune systems.

The discomfort from this social “allergic response,” though at its most extreme and most painful in the United States, is widely shared. Although the social immune systems, like the biological, are essential to health, we are all scratching and sniffing over health care. Over the longer term, however, health care may be an even greater threat to our future capacity to create the wealth which seems so evidently linked with the health of populations.¹⁸ It may stunt our growth.

ENDNOTES

¹*The Economist*, 18 November 1989, 77.

²Eddy van Doorslaer, Adam Wagstaff, and Frans Rutten, eds., *Equity in the Finance and Delivery of Health Care: An International Perspective* (Oxford: Oxford University Press, 1993).

³Edith Rasell, Jared Bernstein, and Kainan Tang, “The Impact of Health Care Financing on Family Budgets,” Economic Policy Institute Briefing Paper (Washington, D.C.: EPI, April 1993).

⁴“It has become common wisdom in both Canada. . .and the United States. . .that a small minority of the population account for a large majority of health care expenditures. This pattern has been found across all age groups. . .[and has] been reported to be consistent through time.” Noralou P. Roos, Evelyn Shapiro, and Robert Tate, “Does a Small Minority of Elderly Account for a Majority of Health Care Expenditures?: A Sixteen-Year Perspective,” *The Milbank Quarterly* 67 (3–4) (1989): 347. Berk and Monheit report that in the United States (in 1987) the highest-using 1 percent of the (noninstitutionalized) population account for 30 percent of all expenditures; the lowest-using 50 percent account for only 3 percent. Per capita expenses for the former are thus on average 500 times those of the latter. Not surprisingly, many more of the high users are the elderly and the poor. Marc L. Berk and Alan C. Monheit, “The Concentration of Health Expenditures: An Update,” *Health Affairs* 11 (4) (Winter 1992): 145–49.

⁵The follow-on impact of an increase in the money supply is another and much more complex matter. But whatever the effects of monetary policy, they do not depend upon whether or not people are also digging holes in the ground.

⁶Steffie Woolhandler and David U. Himmelstein, "The Deteriorating Administrative Efficiency of the U.S. Health Care System," *New England Journal of Medicine* 324 (18) (2 May 1991): 1253–258. This is an estimate of the *excess* cost of overhead in the United States relative to the requirements of a Canadian-style system.

⁷By a remarkably cruel irony, the very fact that relatively inexpensive means are available to protect people against most infectious diseases can itself become a threat to health. Lewis Thomas has labeled as "half-way" technologies the forms of care that can only limit the progress of the illness or its symptoms. The true "high" technologies, by contrast, are those that permit a decisive intervention, to prevent or cure. Lewis Thomas, "The Technology of Medicine," *New England Journal of Medicine* 284 (1971): 1366–368. The iron lung for polio victims exemplifies the former; the polio vaccine the latter. Decisive interventions, like vaccines, can be quite inexpensive. But commercial organizations make money by selling commodities; they cannot produce and sell "health" per se. The most profitable commodities are those embodying the "half-way" technologies, for those with the means to pay for them—again and again and again. Vaccines, on the other hand, "may never prove as enticing as drugs for pharmaceutical makers. They can be quite expensive to develop, and they produce less income—repeated refills not required. Governments and international aid agencies. . . cannot afford pharmacy prices. By one estimate, the annual world-wide vaccine market is about \$2 billion. At \$1.2 billion. . . Prozac. . . might overtake the sale of potions that save the lives of millions of children annually." Gary Stix, "Immuno-Logistics," *Scientific American* 270 (6) (June 1994): 105. Thus one should not be surprised to find a resurgence of infectious disease "in Mogadishu or inner-city Miami." Ibid., 102. If health care is regarded as just another commodity, then only profitable forms of care will, indeed should, be supplied. Health consequences have, from this perspective, no conceptual standing and are irrelevant.

⁸The resurgence of TB in some parts of North America is not a qualification to this success story. It is rather (in varying proportions) a side effect of AIDS, a breakdown of social structure, and a failure of immigration policy.

⁹Aaron Wildavsky, "Doing Better and Feeling Worse: The Political Pathology of Health Policy," *Dædalus* 106 (1) (Winter 1977): 109.

¹⁰United States Congress, *Projections of National Health Expenditures* (Washington, D.C.: Congressional Budget Office, October 1992).

¹¹John R. Saul, *Voltaire's Bastards: The Dictatorship of Reason in the West* (Toronto: Penguin, 1993). Saul discusses the close connections between modern management training and philosophy and the application of "science" to warfare in the modern military officer training schools and staff colleges. His emphasis is less on the self-interestedness of professionalized managers, however, than on their commitment to a particular style of rationality which becomes an end in itself. Saul argues that these "interests," the ideals and world-view of the professional technocrat, are in practice much more dangerous to the rest of us than mere private venality.

¹²Other Americans are also skeptical: ". . . fundamental economic principles. . . put efficient, competitive health care markets in the same class as powdered unicorn

horn." Paul Wallich and Marguerite Holloway, "Health Care without Perverse Incentives," *Scientific American* 260 (1) (July 1993): 109.

¹³Robert G. Evans, Morris L. Barer, and Theodore R. Marmor, eds., *Why Are Some People Healthy and Others Not? The Determinants of Health of Populations* (New York: Aldine De Gruyter, 1994).

¹⁴Or that are merely later broken up for scrap, as in the former Soviet Union.

¹⁵See, for example, Robert W. Fogel, "Nutrition and the Decline in Mortality Since 1700: Some Preliminary Findings," in Stanley L. Engerman and Robert E. Gallman, eds., *Longterm Factors in American Economic Growth* (Chicago, Ill.: University of Chicago Press, 1986), 439-555.

¹⁶There seems to be more general acceptance of the notion that military spending can inhibit growth, although there is still room for debate over both technological spin-offs and Keynesian pump-priming effects. The greater controversy over medical care brings us back to the confusion created by thinking about such care as "goods" in the ordinary sense of the word.

¹⁷Political resistance to both tax increases and program cuts tends to expand government deficits; government borrowing then absorbs savings which might otherwise flow into capital investment. Similarly, increases in the cost of employer-paid benefit programs must either reduce worker incomes, or corporate retained earnings, or payments to shareholders with effects on either corporate or individual savings rates.

¹⁸It may be that it is not only prosperity per se, but (above some minimal level of average income) the distribution of income and of life chances more generally that is most important for health. To the extent that this is so, one might be more concerned with the connections between rates of growth of average incomes and the degree of inequality of incomes at any point in time. In the United States the stagnation of average incomes during the 1980s has been associated with dramatic increases in income inequality, but the causal linkage in either direction is quite another matter.

Can We Have Too Much Health Care?

THE HEALTH-CARE INDUSTRY USES resources, and large amounts of them in most industrialized nations. Recently, health-care expenditures have increased at rates greater than the rates of economic growth in many of these nations. This creates economic and political pressures for governments concerned with sustaining their health-care systems, regardless of the public/private mix in the financing and provision of health care.¹

These tensions are made worse by the strong internal momentum for expansion which is a universal characteristic of the health-care "industry." This, in turn, is fed by several factors, including the attitude of health-care providers who often "take for granted that their work is more important than any other uses a society might have for its resources."²

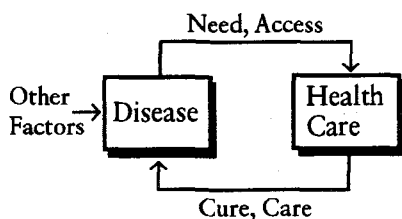
Another important factor, despite accumulated—and accumulating—evidence to the contrary, is the view that health care plays the primary role in determining health.³ *Figure 1* illustrates an oversimplification of this popular view. In this "heating system" model, individuals become sick or injured for a variety of external and unspecified reasons. They present themselves to the health-care system, which professionally interprets their needs for care and responds by delivering the "necessary" services, within whatever institutional arrangements for access and payment exist. These

John N. Lavis is a Researcher at the Centre for Health Economics and Policy Analysis and Department of Clinical Epidemiology and Biostatistics at McMaster University, Canada and a Research Fellow at the Institute for Clinical Evaluative Sciences in Toronto, Canada.

Gregory L. Stoddart is Professor of Health Economics at the Centre for Health Economics and Policy Analysis and Department of Clinical Epidemiology and Biostatistics at McMaster University, Canada.

programs or services reduce the burden of disease in these individuals. The health-care system is both a thermostat and a furnace; the response to colder temperatures (more disease) is to turn up the heat (provide more health care). Health policy effectively becomes health-care policy.

FIGURE 1. A (Too) Simple View of the Determinants of Health

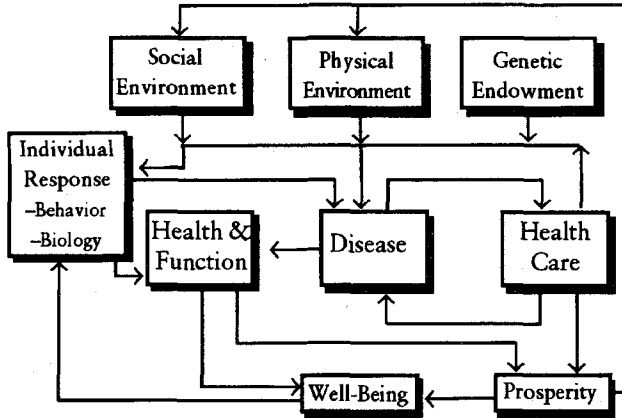


Source: Robert G. Evans and Gregory L. Stoddart, "Producing Health, Consuming Health Care," *Social Science and Medicine* 31 (12) (1990): fig. 1, 1350.

A more appropriate conceptual framework for considering the determinants of health is both more complex and more subtle. One such framework is presented in *Figure 2*. Here an individual's social and physical environment and genetic endowment, along with health care, are identified as the broad determinants of health. The effects of these determinants, both positive and negative, are mediated through the biological and behavioral responses of individuals. For example, an individual's social network (family, friends, and co-workers) may have a positive health effect through biological mechanisms which are at present only partially understood by scientists. This same social network may also have a negative health effect if it fosters the development or continuation of a smoking habit. In this model a distinction is drawn between disease, as recognized and responded to by the health-care system, and health and function, as experienced by an individual. Self-perceived health and function may often be more important to an individual's social performance and well-being. The model also recognizes the exist-

ence of a higher order objective, namely well-being, which is influenced by many factors.

FIGURE 2. A Broader Conceptual Framework for the Determinants of Health



Source: Evans and Stoddart, "Producing Health, Consuming Health Care," fig. 5, 1359.

There are two notable features of this framework. First, *health care* is only one of several determinants of health. And, second, *health* is only one of many determinants of well-being. When health care and health are set in this broader context, it is clear that policies, programs, or services aimed at groups in society or at an entire population may be as important as those aimed at individuals (the usual target of health-care programs or services). Furthermore, policies, programs, or services in sectors other than health care may (as an objective or as a consequence) affect health and well-being.

Industrial safety programs and environmental regulations are two examples of public policies with specific health objectives. These policies are clearly aimed at groups (industrial workers and the general public, respectively). Subsidized housing and day care are examples of public policies developed for reasons other than health but which may have important consequences for health. These policies or programs compete with health care for resources.

For example, the \$350 million budget increase for hospitals in the Canadian province of Ontario in the 1990–1991 fiscal year could have funded seventy thousand more rent-geared-to-income housing units or provided subsidized day care for 547,000 more children.⁴

Figure 2 was developed “to provide meaningful categories in which to insert the various sorts of evidence which are now emerging as to the diverse determinants of health.”⁵ How a problem is framed *is* important. It determines which types of evidence are deemed relevant and which are ignored or discarded. It also determines the set of admissible—or probable—questions for discussion. In this respect, the progression from the “heating system” view of the determinants of health in *Figure 1* to the broader view in *Figure 2*, with its complex interactions and trade-offs, poses a question of increasing importance: Can we have too much health care?

The character of this question could easily be construed as incendiary. However, the intent of this essay is neither to deny the significant contribution that health care makes to health nor to denigrate the dedicated efforts of health-care providers. Rather, the discussion here seeks to provide balance and perspective—a small “voice on the other side”—on the seemingly insatiable appetite of health-care systems for more resources. It is a reminder that health care, like all resource-using activities, has opportunity costs.

The manner in which the question has been posed creates a potential minefield of further questions, definitions, and value judgments. Most important among these are: What is being produced (health care, health, or well-being)? And, against what yardstick (and by whom) is “too much” to be judged? “Health care” is interpreted, in a rather traditional (some may say narrow) way, to be the set of services provided for the primary purpose of restoring, maintaining, or improving health. “Health” is defined as the absence of negative biological circumstances associated with disease or disability. The World Health Organization’s (WHO) definition of health as “a state of complete physical, mental, and social well-being” is so broad as to make all human activity health related and all policy health policy. In this essay, therefore, health, as defined by the WHO, will instead be termed “well-being.”

The yardstick against which production will be judged in this essay is efficiency, including value or worth. The three criteria to be used correspond to the three levels of efficiency in economic theory:

technical efficiency, cost-effectiveness, and allocative efficiency. Production is technically inefficient or cost-ineffective if the same "output" could be produced with, respectively, fewer or less expensive "inputs." Production is allocatively inefficient if an equally valued level or mix of output is possible using fewer resources.

One possible taxonomy of analytically distinct ways in which we can have too much health care is outlined in *Figure 3*. The taxonomy is defined (horizontally) by what is being produced (health care, health, or well-being) and (vertically) by whether production is efficient. Because the production of health or well-being may be approached differently according to whether the focus is on individuals or groups, those columns are subdivided accordingly.

FIGURE 3. A Taxonomy of Ways in Which Too Many Resources can be Devoted to Health Care

		What is being Produced?				
		Health Care	Health		Well-Being	
			Individual	Population	Individual	Population
Is production efficient?	Technically inefficient	2	1		6	
	Cost- ineffective		3			
	Allocatively inefficient	4	5			

1. Health care that is not effective.
2. Effective health care that is more costly than it need be.
3. Health that is more costly than it need be.
4. Health care that is valued at less than its cost.
5. Health that is valued at less than its cost.
6. Well-being that is more costly than it need be.

Although the taxonomy may be very useful as a conceptual sorting device, one should be careful not to draw overly fine distinc-

tions. As the nature of the “product” (health care, health, or well-being) changes, the label of the cell may change, but the fundamental economic problem of how to allocate scarce resources will not change. Furthermore, many of these cells can best be considered together. This leaves us with six ways in which we can have too much (i.e., devote too many resources to) health care if our primary objective is health or well-being.

1. *Health care that is not effective.* If health care is not doing what it is supposed to do (i.e., restore, maintain, or improve health, for an individual or a population) then the resources devoted to it are simply wasted.⁶ In economic terms, there is no “final” output (health), even though an “intermediate” output (health care) has been produced. The failure of health-care programs or services to perform their tasks may occur in several different ways and for different reasons. Some programs or services may actually harm patients or do more harm than good. Others may be ineffective. It may also be the case that an otherwise effective service is rendered ineffective because it is applied in the “wrong” clinical circumstances (i.e., when it is not clinically warranted).

In the Harvard Medical Practice Study adverse events, defined as injuries caused by medical management, “iatrogenesis,” occurred in 3.7 percent of hospitalizations, with 2.6 percent of such events causing permanently disabling injuries and 13.6 percent leading to death.⁷ Ivan Illich has extended the discussion of iatrogenesis beyond clinical iatrogenesis to include the harmful effects of health care at the societal level. He describes social iatrogenesis as the health-denying effects of medicine’s social organization. He posits the medicalization of prevention as a symptom of such iatrogenesis: the responsibility of an individual for his or her future is transformed into the management of that individual by an agency. Illich also describes cultural iatrogenesis, which he defines as the transformation of pain, impairment, and death from a personal challenge to a technical problem, requiring technical solutions from the health-care sector.⁸ Illich’s extensions, although not widely accepted, do suggest the need for a broader understanding of the potentially harmful effects of health care.

As with clinical iatrogenesis, the provision of ineffective health care is rarely an intentional act. Information on effectiveness is

often lacking or not readily available to providers (or is not acted upon when available). A recently published compendium on care in pregnancy and childbirth suggests that of the 282 forms of care evaluated, 125 forms of care require further evaluation and 60 forms of care should be abandoned in light of the available evidence.⁹ Similar systematic reviews of the available evidence on the effectiveness of clinical services have not been undertaken (or completed if undertaken) for other clinical specialties. Even if the information is available, however, the manner in which it is presented to physicians can affect treatment decisions. For example, physicians have been shown to be more likely to treat patients with high blood pressure if the benefits are expressed as a *relative* change in the risk of death (20.3 percent) rather than as an *equivalent* absolute change (from 7.8 percent to 6.3 percent).¹⁰ Factors other than information on effectiveness (such as fear of litigation or method of reimbursement) may also influence treatment decisions. For example, researchers have demonstrated a positive association between the risk of a malpractice claim and the rate of cesarean delivery.¹¹ The effectiveness of health care can also be considered at the level of an entire population. Studies of the association between health-care services and health outcome cast doubt on the effectiveness of health-care services in aggregate.¹²

Inappropriate care is care provided to a patient who does not meet the indications for which a specific intervention has been shown to be effective. Inappropriate provision has been found to range from 4 to 27 percent for coronary angiography, from 2 to 16 percent for coronary artery bypass surgery, from 13 to 32 percent for carotid endarterectomy, and from 11 to 24 percent for gastrointestinal endoscopy.¹³ Hence, a not insignificant number of otherwise effective services are provided to patients who will not benefit from them.

2. *Effective health care that is more costly than it need be.* Health care becomes "costly" when the same quantity of services could be obtained with a different technical or physical deployment of fewer inputs (personnel, equipment, facilities, or know-how). By the same reasoning, a different deployment or combination of the resources already committed to health care could result in a higher level of output (i.e., more services) with the same number of inputs. For

example, because of duplication of administrative departments and personnel, US hospital administrative costs are more than twice those in Canada.¹⁴ Also, hospitals may have low occupancy rates and laboratories may have long periods of downtime. Health policy reforms which reduce the costs of bureaucracy or reduce excess capacity could yield substantial savings (or permit the production of a higher level of output).

"Costly" may also be defined in a monetary sense, taking into account the prices of the inputs used in the production of health care. In this case, health care is considered costly if the same quantity of services could be obtained with a lower level of expenditure. In other words, although the services may be effective, they are not cost-effective. For example, health-care provision is more costly than it need be if the same services could be provided by a less expensive but equally effective health-care provider. If nurse practitioners were utilized in the Canadian health-care system as fully as possible (given existing evidence of their safety and effectiveness), it has been estimated that savings would accrue in the range of 10 to 15 percent for medical services as a whole and 16 to 24 percent for ambulatory services.¹⁵ The provision of health care may also be more costly than necessary if it is provided in a more capital-intensive setting than needed. Estimates of the extent of nonacute hospital days of care in the United States range from 20 to 48 percent for adults and 5 to 21 percent for children.¹⁶ These days of care could have been provided in nursing homes, ambulatory care facilities, or other less capital-intensive settings.

3. *Health that is more costly than it need be.* Health care is only an intermediate output. Resources are committed to it for its ability to contribute to another production process, the output of which is health. Just as health care can be produced in more costly or less costly ways (in either a technical or monetary sense) so, too, is it possible to produce health in more costly ways than necessary. Once it is accepted that health care is but one of the determinants of health, it becomes apparent that the phrase "too many resources devoted to health care" could also refer to the circumstance in which a specific level of health could be achieved by reducing expenditures on some types of health care relative to others, or by spending less on health care and more on other contributors to

health. Consider, for example, a fixed budget of \$1 million and the years of life which could be bought with it. Spending the money on monitoring low-risk patients in coronary care units would buy three years; providing propranolol for mild to moderate high blood pressure would buy seventy-three years; and providing bypass surgery for left-main coronary artery disease in middle-aged men would buy 134 years.¹⁷ A similar exercise could be performed using such measures of health as improvements in functional status or improvements in the quality of life. However, it is important to consider who gets these years or improvements in functional status and quality of life; this type of analysis does not typically address this concern.¹⁸

Even cost-effective health care can be seen as competing for resources with other activities that also contribute to health.¹⁹ Among the few examples of empirical research that compare health care with health-producing strategies outside the health-care system one notable example is a comparison of the cost-effectiveness of six strategies to reduce infant mortality: teenage family planning; a supplemental food program for women, infants, and children; community health centers with maternal and infant care projects; abortion; prenatal care; and neonatal intensive care. The early initiation of prenatal care was found to be the most cost-effective strategy and neonatal care one of the least cost-effective; supplemental food programs consistently ranked among the most cost-effective.²⁰

The need for further work on health-producing strategies beyond the formal health-care system is suggested by international comparisons of total health expenditures and population health indicators. For example, in 1991, Japan, Canada, and the United States devoted 6.8, 10.0, and 13.2 percent, respectively, of their gross domestic product (GDP) to the health-care sector.²¹ The pattern of infant mortality rates is the reverse of what one would expect based on the expenditure data given above: 4.6, 6.8, and 9.1 per 1,000 live births in Japan, Canada, and the United States, respectively. Although this is far from definitive evidence, some countries (such as Japan) appear able to produce "more" health with a lower level of health-care sector resource commitment.

4. *Health care that is valued at less than its cost.* Thus far, interpretations of "too much" have been based solely on produc-

tion (supply-side) considerations. The production of health care, or the health that it produces in conjunction with other determinants, has been more costly than necessary for one reason or another. But cost and supply constitute only one side of an economic ledger. Just because something is, or can be, produced cost-effectively does not mean that it should be produced at all or at a particular level or rate. That judgment depends on an assessment of benefits as well as costs (i.e., "worth") and brings in consumption (demand-side) considerations. In other words, the benefits of health care must be valued in relation to its costs.

This valuation process turns out to be much more difficult for health care than for other goods and services. Prices, which normally link the resource cost of a good or service with its value to consumers, typically do not exist in health-care markets because of insurance and/or extensive public financing. Also, the criteria for judging whether different allocations of health care are "better" or "worse" are complex and controversial.²² Despite these practical difficulties, it does not seem unreasonable to suggest that if the benefits of health care are perceived²³ by those receiving or paying for care to be less than the costs of care (i.e., other things foregone), then there is "too much" care being produced and consumed. This may mean that too many resources are devoted to health care in aggregate or, more likely, that too many resources are devoted to certain types of care, that the *mix* of health-care outputs is wrong.

At the individual level, patients' and providers' valuations of benefits in relation to costs may differ. In addition, patients' own valuations may differ according to how the information is presented to them. For example, male patients' perceptions of how they are bothered by the symptoms of benign prostate gland enlargement and their willingness to undertake the risk of operation appear to differ from their physicians' perceptions. Preliminary results have shown that among patients who used interactive videodiscs that explained treatment options and associated risks, 25 percent of men who preferred surgical removal of the prostate gland after discussion with their physician changed their minds and opted for "watchful waiting."²⁴ Research on patient preferences for lung cancer therapies has shown that the attractiveness of surgery, relative to radiation therapy, was substantially greater when the problem was framed in terms of the likelihood of living rather than dying.²⁵

The benefits of health care may also be valued at less than its cost where groups are concerned. Oregon's effort to extend health-care coverage within a limited budget provides an example of an attempt to solicit and apply citizens' values systematically in a prioritization process.²⁶ The result of this process was a recommendation that some health-care services be removed from insurance coverage.

5. *Health that is valued at less than its cost.* The primary benefit of health care and the reason that it is of value is the contribution that it makes to health. But health is itself subject to the same criterion of "worth." It may be that the health produced by health care and other resource-consuming determinants is not valued as highly as other outputs that could have been produced with the same resources. In other words, the mix of health versus other things (such as consumer goods, education, security, and justice) or the mix of different types of health could be improved.

Advance directives provide one dramatic example of the valuation of health relative to its cost. Such directives specify the medical care which an individual wishes to receive (or not receive) if he or she is no longer able to make that decision. A survey of outpatients of primary care physicians found that most patients desired advance directives and would refuse life-sustaining treatments in the event of a persistent vegetative state or dementia with a terminal illness.²⁷ These health states were valued less than the (nonmonetary) costs associated with life-sustaining treatment.

Less dramatic examples of the valuation of health occur frequently in public sector decision-making. When local, regional, or national governments decide, for example, to improve the quality of education by increasing teacher-to-student ratios instead of reducing the burden of communicable diseases through a universal immunization program, they are implicitly, if not explicitly, valuing improvement in education more than improvement in health.

6. *Well-being that is more costly than it need be.* Just as health care is but one determinant of health, so too is health only one determinant of well-being.²⁸ It is therefore possible, in principle, to produce well-being in more costly or less costly ways, depending upon the amounts of health and other inputs that are used and the respective costs of the inputs. Too many resources can be devoted to health care if the level of health which results is too high—in

other words if less health or more of other contributors would increase well-being. Well-being is the goal toward which health and its competitors for resources make their ultimate contribution, and against which the worth of the contributions is judged. This higher order concept is indeed the “end of the line,” for it is presumably not possible to have too much well-being!

Yet, in health policy debates in industrialized countries during the past several decades the “end of the line” has been health care and hence the focus of debate has been the means and cost of its production. The broad view of the determinants of health presented in *Figure 2* allowed a very different question to be explored in this essay. The end of the line was defined as health (as an input to well-being) or well-being, and hence the focus of inquiry was the means and cost of *their* production.

The question of whether we *can* devote too many resources to health care is, of course, quite different from the question of whether we *do* devote too many resources to health care. The taxonomy and supporting examples used in this essay have demonstrated that it is possible to have too much health care. Whether or not we *do* devote too many resources to health care depends on social judgments about what constitutes well-being, social judgments about the equity of reallocations involving less health care, and the strength of the relevant evidence which exists for any particular cell depicted in *Figure 3*.

Measurements of “health,” such as life expectancy, are well-established and routinely available. The measurement of well-being hinges on a social judgment of what constitutes well-being. Economists often view the per capita GDP of a country as a measure of “welfare”—a concept analogous to well-being. But an index of production is a woefully incomplete measure. There are clearly other characteristics relevant to social well-being: life expectancy,²⁹ literacy, income equality, and personal safety, to name but a few. The relative contribution of these determinants of well-being may vary across groups and over time. Nonetheless, a social judgment of what constitutes well-being is needed if well-being is to be one of the objectives against which health-care expenditures are judged.

The fairness, or “equity” as it is often termed by policy analysts, of alternative resource allocations is also a matter which requires

social judgments. The taxonomy in *Figure 3* masks these complex issues of equity in the distribution of both the benefits and the costs of health care and other activities. Aggregation across a population is not simple either in theory or practice because individuals are not identical in their resources, needs, or preferences. Policies which may "improve" performance in the sense that they remove inefficiencies in some or all of the cells of *Figure 3* are not unequivocally "better." That judgment must be left to society.

Even with social judgments regarding the constituent elements of well-being and the equity of health-care reallocations, responses to the question of whether we *do* devote too many resources to health care are constrained by a lack of information regarding the production of health and well-being. The availability and strength of the evidence needed to inform this debate varies widely across the cells in the taxonomy. A reasonably extensive empirical literature suggests that we *do* in fact devote too many resources to health care in certain ways (such as health care that is not effective or effective health care that is more costly than it need be). But there is little empirical research on either the complex trade-offs between different mixes of inputs involving less health care and more of "other things" which contribute to health, or less health and more of "other things" which contribute to well-being.

The opportunity cost of health-care spending is the value of the "other things" which could have been purchased with available resources. The opportunities foregone because of health-care spending include activities in sectors other than health care which may (as an objective or a consequence) affect health and well-being. Perhaps the most thought-provoking foregone opportunity is the opportunity to commit more resources to activities producing economic growth, which in turn could both augment the capacity of the economy to sustain future health-care spending *and* improve health and well-being (both directly and indirectly through improvements in health). Exactly *how* health-care spending affects economic growth is a subject with more questions than answers.

It can be argued that health-care spending represents an "investment" in health which increases the capacity of individuals to produce goods and services. However, the increasing concentration of health care on persons outside the labor force, such as the chronically ill or the very old, weakens this argument.³⁰ Another

positive effect of health-care spending, it is sometimes argued, is that it creates jobs. But many would argue that if the jobs created are those traditionally associated with the health-care sector—health-care providers and their support staff—such spending represents redistribution rather than growth.

However, it can also be said that “other activities” such as savings and investment, which may be reduced by health-care spending, would have otherwise found their way into research and development or other activities which contribute to economic growth. Furthermore, the health-care sector employs many talented individuals who could make substantial contributions to these activities. In this view, health-care spending has a very high opportunity cost: it threatens the engine of growth which will provide the future capacity to sustain the social infrastructure (including the health-care system) essential to health and well-being.

Even the World Competitiveness Report, produced by the World Economic Forum, has oscillated on the question of whether health-care spending has a positive or negative net effect on competitiveness (and therefore economic growth). This much-quoted publication attempts, on an annual basis, to determine the extent to which “national environments [are] conducive or detrimental to the domestic and global competitiveness of enterprises operating in those countries.” In 1991, Canada was ranked first on the criterion related to health care: quality and availability of health care to employees and their families. The view taken by the Report was that “high spending is conducive to competitiveness as it keeps the work force healthy and productive.”³¹ In 1992, however, Canada was ranked eighteenth on criteria related to health: central government expenditures on health, population per physician, and population per nurse.³² No explanation was given for this sudden reevaluation but clearly the Forum’s view of health-care spending had changed.

The link between health-care spending and economic growth is only one example of a trade-off involving less health care and more of “other things” which contribute to health or well-being. For the evaluation of health-producing strategies beyond the formal health-care system, intersectoral economic evaluation methods need to be developed and applied. This would enable a comparison of the costs of alternative strategies (such as job retraining for the unem-

ployed versus drug therapy for depression) against a common measure of effect (such as improvements in health-related quality of life or life expectancy). These trade-offs will seldom be made explicit if health care remains at center stage and health policy remains, essentially, health-care policy. The possibility of redirecting resources from the health-care sector to subsidized day care or to investments for economic growth, for example, receives little attention. Yet these changes may have a more profound impact on the health and well-being of individuals and populations than could be accomplished through the health-care system. When applied narrowly to health care, the popular wisdom is wrong: it is possible to have too much of a good thing.

ACKNOWLEDGMENTS

We wish to thank Jeremiah Hurley, Jonathan Lomas, the Polinomies Group at McMaster University, and participants at the Honda Foundation Discoveries Symposium for stimulating comments on an earlier version of this paper. We take responsibility for remaining errors or omissions.

ENDNOTES

¹There is no purely private system anywhere in the industrialized world, although models of publicly financed and (largely) privately provided health care exist (such as the Canadian system). The percentage of public financing varies considerably: from 36 percent and 42 percent in Turkey and the United States, respectively, to 90 percent or more in Sweden, Luxembourg, and Norway in 1990; Canada and Japan are somewhere in the middle of the OECD pack, at 73 percent and 72 percent, respectively. See George J. Schieber et al., "U.S. Health Expenditure Performance: An International Comparison and Data Update," *Health Care Financing Review* 13 (4) (1992): 1.

²Robert G. Evans, "Health Care Reform: 'The Issue from Hell'," *Policy Options* (July-August 1993): 38.

³Robert G. Evans and Gregory L. Stoddart, "Producing Health, Consuming Health Care," *Social Science and Medicine* 31 (12) (1990): 1347-363 passim.

⁴Ronald Labonte, "Health Care Spending as a Risk to Health," *Canadian Journal of Public Health* (July/August 1990): 251-52.

⁵Evans and Stoddart, "Producing Health, Consuming Health Care," 1349.

⁶See Jan Blustein and Theodore R. Marmor, "Cutting Waste By Making Rules: Promises, Pitfalls, and Realistic Prospects," *University of Pennsylvania Law Re-*

view 140 (5) (May 1992): 1543–572, for a taxonomy of waste that includes some of the elements in the broader taxonomy developed in this paper.

⁷Troyen A. Brennan et al., “Incidence of Adverse Events and Negligence in Hospitalized Patients,” *New England Journal of Medicine* 324 (6) (7 February 1991): 370–76.

⁸Ivan Illich, *Limits to Medicine* (London: Penguin Books, 1976).

⁹Iain Chalmers et al., *Effective Care in Pregnancy and Childbirth* (Oxford: Oxford University Press, 1989); Murray Enkin et al., *A Guide to Effective Care in Pregnancy and Childbirth* (Oxford: Oxford University Press, 1989).

¹⁰Lachlan Forrow et al., “Absolutely Relative: How Research Results Are Summarized Can Affect Treatment Decisions,” *The American Journal of Medicine* 92 (February 1992): 121–24.

¹¹A. Russell Localio et al., “Relationship Between Malpractice Claims and Cesarean Delivery,” *JAMA* 269 (3) (1993): 366–73.

¹²For examples of studies which examine the effectiveness of health care at the level of a population, see John B. McKinlay et al., “A Review of the Evidence Concerning the Impact of Medical Measures on Recent Mortality and Morbidity in the United States,” *International Journal of Health Services* 19 (2) (1989): 181–208; and John M. Neutze and Harvey D. White, “What Contribution Has Cardiac Surgery Made to the Decline in Mortality from Coronary Heart Disease,” *British Medical Journal* 294 (14 February 1987): 405–409.

¹³For a summary of this literature, see John N. Lavis and Geoffrey Anderson, “Inappropriate Hospital Use in Canada: Definition, Measurement, Determinants and Policy Implications,” *Cost-Effectiveness of the Canadian Health Care System*, Queen’s-University of Ottawa Economic Projects, University of Ottawa Graduate School Working Paper 93–08.

¹⁴Steffie Woolhandler et al., “Administrative Costs in U.S. Hospitals,” *New England Journal of Medicine* 329 (6) (5 August 1993): 400–403.

¹⁵Frank T. Denton et al., “Potential Savings from the Adoption of Nurse Practitioner Technology in the Canadian Health Care System,” *Socio-Economic Planning Sciences* 17 (4) (1983): 199–209.

¹⁶Lavis and Anderson, “Inappropriate Hospital Use in Canada: Definition, Measurement, Determinants and Policy Implications.”

¹⁷Louise B. Russell, “Opportunity Costs in Modern Medicine,” *Health Affairs* (Summer 1992): 164.

¹⁸For a discussion of several difficulties associated with the use of “league tables” of the cost-effectiveness of alternative interventions, see Michael Drummond et al., “Cost-Effectiveness League Tables: More Harm Than Good?,” *Social Science and Medicine* 37 (1) (1993): 33–40.

¹⁹Of course, if health care is not effective, as discussed above, it cannot be cost-effective, and cost-effectiveness in the production of health care is but one step on the road to cost-effectiveness in the production of health.

- ²⁰Theodore Joyce et al., "A Cost-Effectiveness Analysis of Strategies to Reduce Infant Mortality," *Medical Care* 26 (4) (April 1988): 348–60.
- ²¹George J. Schieber et al., "Health Spending, Delivery, and Outcomes in OECD Countries," *Health Affairs* (Summer 1993): 120–29.
- ²²See Anthony J. Culyer, "The Normative Economics of Health Care Finance and Provision," *Oxford Review of Economic Policy* 5 (1): 34–58; and Uwe E. Reinhardt, "Reflections on the Meaning of Efficiency: Can Efficiency Be Separated from Equity," *Yale Law and Policy Review* 10 (302) (1992): 302–315.
- ²³Or would be so perceived if the recipients/payers were fully informed. Another way of thinking of such "allocative efficiency" might be to imagine what fully informed individuals or societies might wish to "spend" on health care, and compare it to the current pattern.
- ²⁴John E. Wennberg, "Better Policy to Promote the Evaluative Clinical Sciences," *Quality Assurance in Health Care* 2 (1): 21–29; Jeff Brooke, "Educational Video-discs Give Patients Information to Make the Right Choice," *Medical Post*, 16 March 1993. Although several studies suggest that fully-informed patients choose less health care, it is not clear that enhancing the role of patients in the decision-making process will always lead to less health care. See, for example, Mark N. Levine et al., "A Bedside Decision Instrument to Elicit a Patient's Preference Concerning Adjuvant Chemotherapy for Breast Cancer," *Annals of Internal Medicine* 117 (1) (1 July 1992): 53–58.
- ²⁵Barbara J. McNeil et al., "On the Elicitation of Preferences for Alternative Therapies," *New England Journal of Medicine* 306 (21) (27 May 1982): 1259–262. In addition to ensuring that the benefits of health care are perceived by those receiving care as greater than the costs, the involvement of patients in decisions about their care has been shown to improve patient outcomes. Sheldon Greenfield et al., "Expanding Patient Involvement in Care: Effects on Patient Outcomes," *Annals of Internal Medicine* 102 (1985): 520–28.
- ²⁶For a description of the process used, see Daniel M. Fox and Howard M. Leichter, "Rationing Care in Oregon: The New Accountability," *Health Affairs* (Summer 1991): 21.
- ²⁷Linda L. Emanuel et al., "Advance Directives for Medical Care—A Case for Greater Use," *New England Journal of Medicine* 324 (13) (28 March 1991): 889–95.
- ²⁸Jonathan Lomas, in his remarks as a discussant at the Honda Foundation Discoveries Symposium, pointed out that health care may be a determinant of well-being while leaving health unaffected. "At the level of an individual. . . the placebo effect is based upon generating perceptions of benefit in the absence of demonstrable gains in health. . . ." At the level of a broader social grouping, such as a community, "attempts to close a [community] hospital, under the judgment that it is not 'needed' to maintain health, come up hard against the perception of the community that it is needed as a major contributor to social cohesion, local identity and, ironically, the community's capacity to attract new industry for economic growth. The latter are likely significant contributors to well-being."
- ²⁹Amartya Sen, "The Economics of Life and Death," *Scientific American* (May 1993): 40–47.

³⁰Evans and Stoddart, "Producing Health, Consuming Health Care," 1359. This does not mean, however, that health care should not be provided for these groups. Rather, as the authors note, its "benefits must be found in the value of the resulting improvements in health, not in some further productivity gains."

³¹World Economic Forum, *World Competitiveness Report 1991*, 11th ed. (Geneva: World Economic Forum, 1991), 8, 39, 206.

³²World Economic Forum, *World Competitiveness Report 1992*, 12th ed. (Geneva: World Economic Forum, 1992), 45, 652-54.

The Epidemiological Transition: From Material Scarcity to Social Disadvantage?

INTRODUCTION

HEALTH AND WEALTH HAVE ALWAYS appeared to be closely related. But within that relationship there is an important historical discontinuity which not only tells us about the changing determinants of health, but also marks a fundamental change in the limiting constraints on the quality of life in modern societies. Mortality rates in the developed world are no longer related to per capita economic growth, but are related instead to the scale of income inequality in each society. This represents a transition from the primacy of material constraints to social constraints as the limiting condition on the quality of human life.¹

The implications of this change are not confined to health. Health can be viewed as a general indicator of welfare and the effects of social and economic change. If health is influenced more by income distribution than by economic growth, then the same is likely to be true of other aspects of human welfare. Furthermore, if increases in the quality of life now depend primarily on improving the social fabric of society rather than on general rises in prosperity, then we must ask whether further *undifferentiated* economic growth is worth the environmental risks.

LIFE EXPECTANCY AND GROSS NATIONAL PRODUCT PER CAPITA

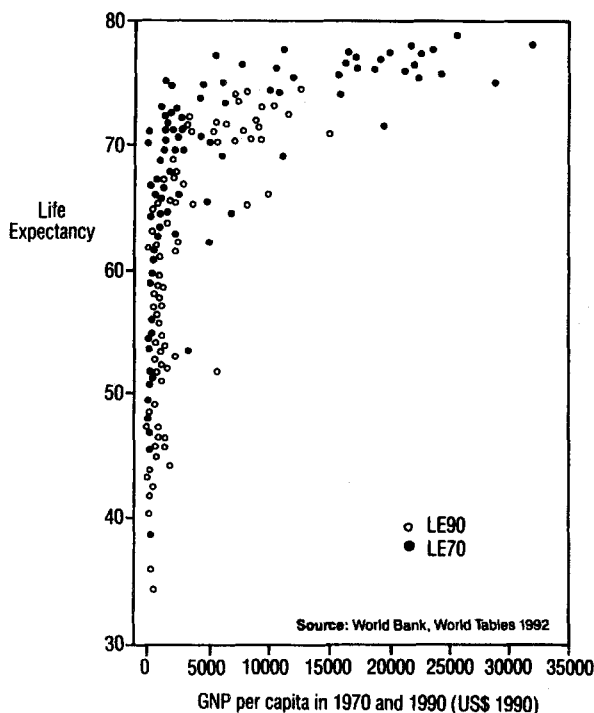
Rising living standards were the basis of the historical decline of mortality in the developed world. While it is possible to argue the

Richard G. Wilkinson is Senior Research Fellow at The Trafford Centre for Medical Research, University of Sussex, Brighton, United Kingdom.

relative historical contributions of better nutrition, sewers, clean water supplies, improved housing, and, eventually, immunization to the long decline in mortality rates in the developed world, there can be no doubt that the enabling and sustaining power of economic growth was behind them all.

Evidence of the broad relationship between economic growth and increasing life expectancy can be found not only in the history of the developed countries, but also in the Third World today. *Figure 1* shows the relationship between Gross National Product per capita (GNPpc) and life expectancy at birth in rich and poor countries in 1970 and in 1990.

FIGURE 1. Life Expectancy at Birth (Male and Female) in Relation to GNP per capita, 1970 and 1990



Although it is clear that life expectancy rises steeply with increased GNPpc among the poorer countries, the data in *Figure 1* also suggests that the relationship between GNPpc and mortality

peters out in the developed world. Apparently, there is some minimum level of income (around \$5,000 per capita in 1990) above which the absolute standard of living ceases to have much impact on health.

Putting income on a log scale, as in Figure 2, it appears more accurate to think in terms of radically diminishing health returns to increasing income. However, as the curve in Figure 2 indicates, as countries get richer even equally large *proportional* changes in incomes produce diminishing absolute returns in life expectancy.

FIGURE 2. Life Expectancy and GDP per capita, 1990

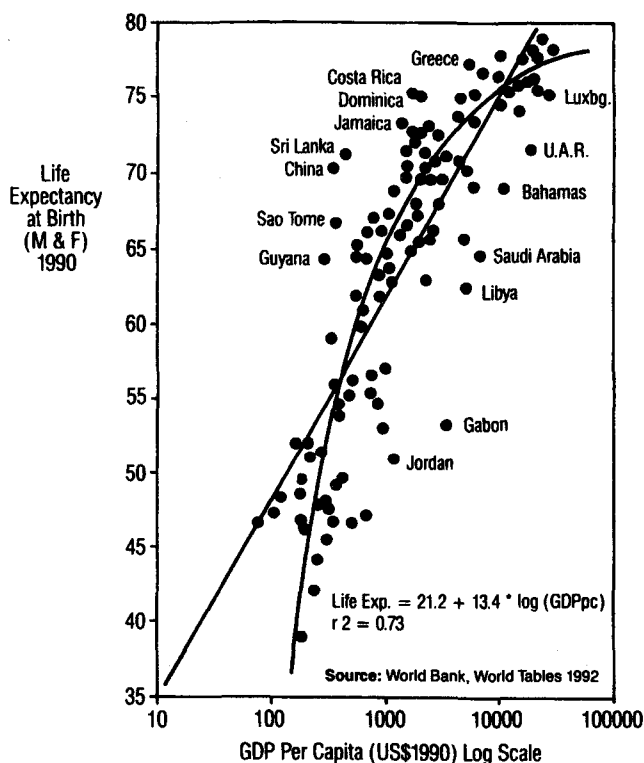
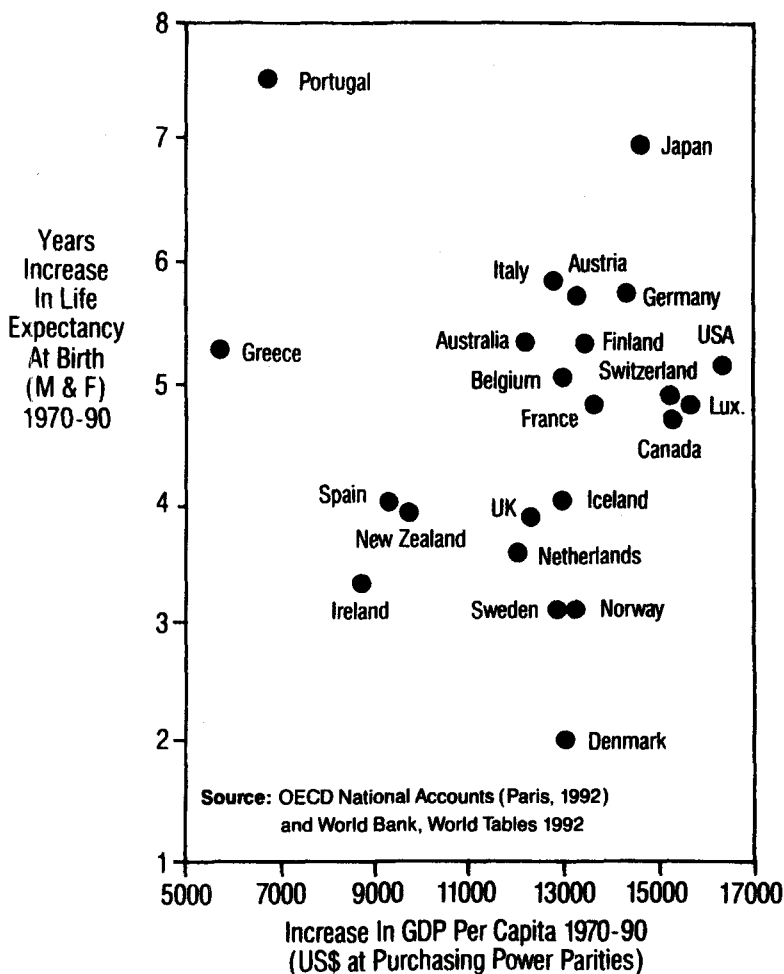


Figure 3 shows that between 1970 and 1990 there was very little correspondence between changes in the Gross Domestic Product per capita (GDPpc) and life expectancy in the developed countries. In this example, GDPpc is converted at purchasing power parities

to provide more accurate comparisons of the real material standard of living.

FIGURE 3. Increases in Life Expectancy Related to Increases in GDP per capita in OECD Countries, 1970–1990



Figures 1, 2, and 3 show that regardless of whether one looks at Organization for Economic Cooperation and Development (OECD) countries at a single point in time or over several decades, there is no strong relationship between income and health among the devel-

oped countries.² People in one country can be twice as well-off on average as those in another without benefit to their mortality rates. As *Figure 3* shows, one country's economy can grow twice as fast as another's for twenty years without having an effect on life expectancy.

APPROACHING LIMITS OF HUMAN LIFE EXPECTANCY?

One obvious explanation for the virtual disappearance of a relationship between GNPpc and life expectancy is that it reflects the approaching limits of human life expectancy. However, this fails to explain why the rate of improvement in life expectancy is not slowing down where longevity is already high. Life expectancy continues to increase at 2 to 2.5 years per decade among the rich countries, as it has during most of this century (wartime excluded). In addition, it appears that in several developed countries the fastest improvements are coming from those sections of the population where life expectancy is highest. For these reasons, we must assume that the leveling of the curve of rising life expectancy with GNPpc (shown in *Figure 1*) indicates the declining health benefits of further increases in the standard of living in the developed world.

THE EPIDEMIOLOGICAL TRANSITION

The "epidemiological transition" is the shift in the main causes of death—from infectious diseases to degenerative cardiovascular diseases and cancers. It marks a fundamental change in the main determinants of health and seems to indicate the point in economic development at which the vast majority of the population gained reliable access to the basic material necessities of life.

The impact of medical science is not reflected in the epidemiological transition. In fact, the transition would have happened (and largely did happen) without it. The great infectious diseases of the nineteenth and early twentieth centuries dwindled to a fraction of what they had been long before immunization or effective medical treatment became available. Strongly associated with poverty, in the past as in the Third World today, the decline of the great infections reflected improved living standards and conditions.

The effect of economic development is also illustrated by the way in which the epidemiological transition saw the so-called "diseases of affluence" transformed into the diseases of the poor. No longer primarily "businessmen's diseases," coronary heart disease, stroke, and obesity have all become more common among the less well-off.³ That this transition was associated with the attainment of a certain level of material wealth is shown by the associated reversal in the social gradient in smoking and in the consumption of refined foods such as white bread and sugar. In all but a few countries the evidence on cardiovascular diseases is fragmentary, but it seems likely that many developed countries went through the same transition and experienced the same change in the class distribution of these diseases, although at different times.

The change in the social class distribution of obesity is much more widely recognized. Throughout most of history obesity was a condition which affected only the privileged, and in many cultures was a status symbol—embodied proof of prosperity. The rich were fat and the poor were thin. This remained true in the developed world well into the present century when rising living standards began to enable the poor to be fat, and is still true in some parts of the Third World.

In Britain the change appears to have been accelerated by nutritional changes during World War II.⁴ The Chief Medical Officer in the British Ministry of Education commented on the increasing incidence of obesity in schoolchildren in his annual reports of 1956–1957 and 1960–1961.⁵ As if to confirm Pierre Bourdieu's thesis of the importance of the desire to express social distinction in aesthetic judgments, as soon as the poor became able to be fat, obesity ceased to be a status symbol.⁶

A further indication of the widespread attainment of adequate material standards for the populations of developed countries comes from trends in the proportion of babies with birth weights below 2,500 grams. Although perinatal and infant mortality rates have continued to decline rapidly, since the 1950s in England and Wales there has been no further marked decline in the proportion of low birth weight babies: between 6 and 7 percent of babies have weighed less than 2,500 grams ever since.⁷ This proportion of low birth weight babies seems to represent a residual core which is unaffected by improvements in average prosperity.

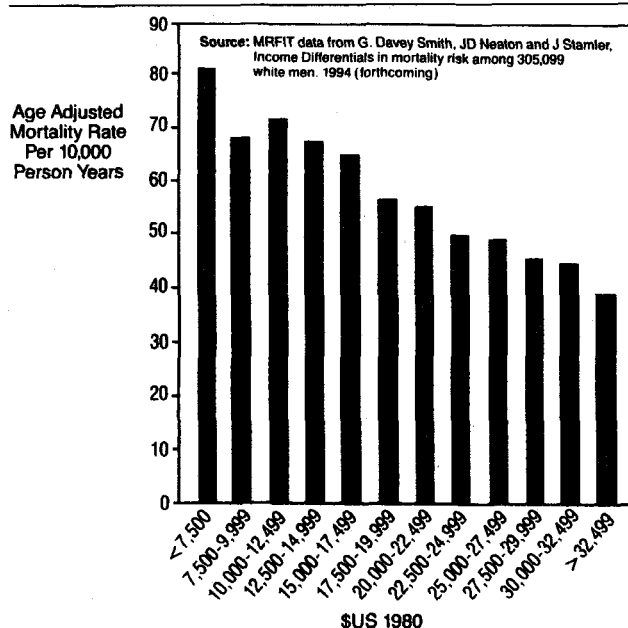
Thus, it seems that in the later stages of industrial development countries go through a health climacteric after which the health of the vast majority of the population is no longer substantially affected by the absolute material standard of living.

INCOME DISTRIBUTION

After the epidemiological transition, health remains closely associated with deprivation, but the relationship now is with relative rather than absolute deprivation. Health within developed countries continues to show a clear gradient with measures of socioeconomic status.⁸ Figure 4 shows how death rates vary according to income group in the United States. Evidence from Britain, covering most of the present century, suggests that the social class differences in death rates have increased or decreased as the proportion of the population living in *relative* poverty has increased or decreased.⁹

There is, however, a marked contrast between the well-ordered relationship between socioeconomic status and health within a country (as shown in Figure 4) and the lack of any clear relationship between socioeconomic status and health between countries (as shown in Figures 1 and 3). This contrast is extremely important.

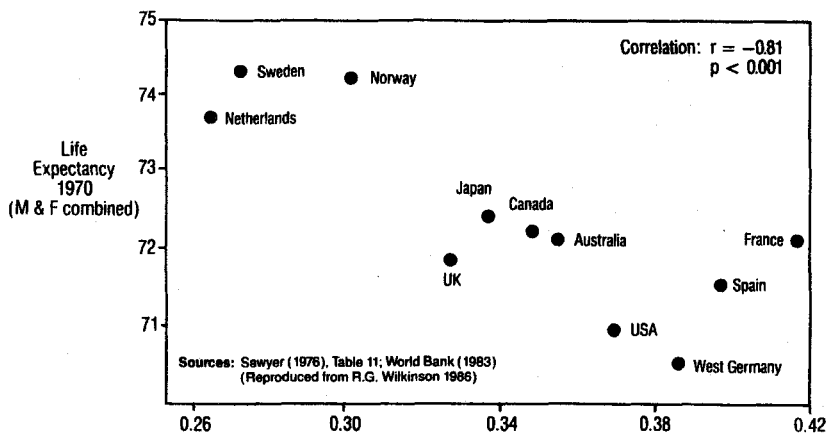
FIGURE 4. Income and Mortality among Three Hundred Thousand White US Men



Given that the points in *Figures 1 and 3* are whole nations, the fact that they do not even begin to line up like those for income groups within the United States cannot be chalked up to sampling error or the like. Something quite different is going on. Health is affected by differences in *relative* income (differences between groups of people within the same society), not by the absolute level of average incomes for each society as a whole. This is confirmed by the surprisingly strong relationship between income distribution and mortality rates in developed countries.¹⁰

The countries with the longest life expectancy are not the wealthiest but those with the smallest spread of incomes and the smallest proportion of the population in relative poverty. Using a standard measure of income distribution, *Figure 5* shows that in 1970 the countries with the smallest income differences were also the countries where average life expectancy was longest. Since then the rate of growth of average life expectancy has been closely related to changes in income distribution. *Figure 6* shows that in the twelve member states of the European Community (EC), average life expectancy between 1975 and 1985 has grown fastest in those countries where relative poverty decreased fastest (or increased slowest).

FIGURE 5. Life Expectancy and Income Distribution in OECD Countries

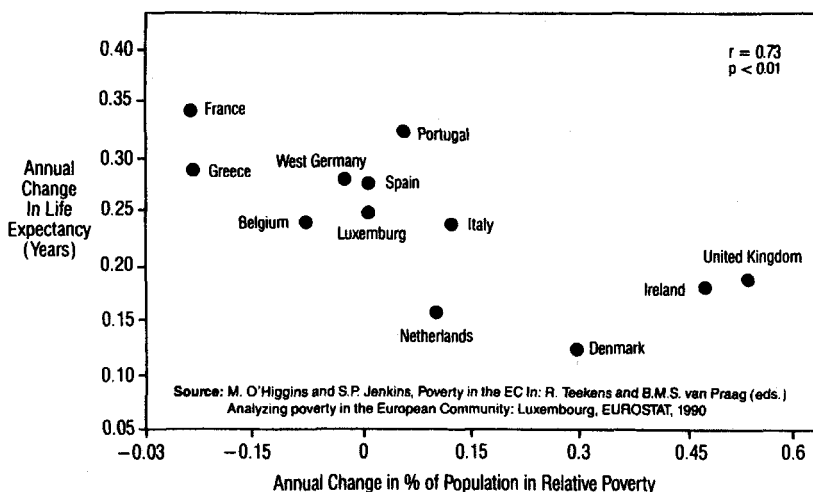


Gini coefficients of post-tax income inequality, standardized for household size (inequality increases to the right)

The relationship implies that between one-half and three-quarters of the differences in average life expectancy from one developed country to another may be attributed solely to differences in income distribution. Controlling for GNPpc, public expenditure on medical care, or total public expenditure, does not suggest that the relationship with income distribution is a proxy for an effect of better public services. Nor are such factors thought to have a sufficiently powerful influence on national mortality rates to make them potential candidates.

Because this relationship is established on national data it may prove to be an ecological fallacy. However, there is substantial evidence not only that individual health and income are related (as shown in *Figure 4*), but also that health is responsive to changes in income.¹¹ The function of the international evidence shown here is primarily to distinguish between relative and absolute income levels. As relative income is inherently a social concept it cannot be dealt with at an individual level: societies, not individuals, have income distributions.

FIGURE 6. The Annual Rate of Change of Life Expectancy in Twelve EC Countries, 1975–1985 (with the rate of change in the percentage of the population in relative poverty)



A telling example of the influence of income distribution on health is the contrast between health trends in Britain and Japan.

Since 1970 income distribution in Japan has narrowed significantly, making it the narrowest of any country reporting to the World Bank. At the same time, life expectancy in Japan has increased by 6.9 years and is now the highest in the world. In contrast, income distribution in Britain has widened, and Britain's position on the life expectancy table of OECD countries has fallen from tenth in 1970 to seventeenth in 1990, with a gain of only 3.9 years in life expectancy.

As income differences in Japan have narrowed, so have differences in death rates among social classes.¹² In Britain, on the other hand, both income and mortality differences have widened.¹³ These divergent trends lie behind the contrasting gains in average life expectancy. The likelihood that there is a causal relationship between income distribution and national life expectancy is also confirmed by recent evidence of the relationship between inequalities in both health and income in different countries.¹⁴

PSYCHOSOCIAL INTERPRETATION

The fact that health seems to be influenced more by differences in income than by average level of income suggests that cognitive processes of social comparison are involved. The importance of relative income to health suggests that psychosocial factors related to deprivation and disadvantage are involved. That is to say, it is less a matter of the immediate physical effects of inferior material conditions than of the social meanings attached to those conditions and how people feel about their circumstances and about themselves.

The need to develop a psychosocial understanding of the impact of material differences on health is indicated by three other aspects of the relationship between health and socioeconomic status. 1) The socioeconomic gradient in health does not distinguish merely between the poor and the rest of society, but affects every rung of the social ladder. Thus, people who own houses and have two cars are healthier than those who rent houses and have one car;¹⁵ administrative civil servants are healthier than executives;¹⁶ and people in the highest income group are healthier than those only slightly less well-off (see *Figure 4*). 2) Despite rising real incomes, which have undoubtedly reduced the incidence during the second half of this

century of absolute material deprivation, the mortality disadvantage of blue collar workers and their families appears to have increased in several countries. 3) In Britain, now one of the poorer developed countries, absolute living standards among the *poorest* 20 percent of the population are remarkably high. In 1992, some 72 percent had central heating, almost all had televisions and refrigerators, 72 percent had telephones, and almost 60 percent had videocassette recorders.¹⁷

The socioeconomic gradient in mortality (*Figure 4*) could only be attributed to the direct effects of material factors if it is assumed that increasing standards of comfort higher up the income scale have as great an effect as material improvements further down the scale. However, if this were so, there would be a closer association between changes in GDPpc and mortality than is shown in *Figure 3*.

There are a number of different ways in which psychosocial links might be involved in the association between socioeconomic status and health. For example, the stresses of economic insecurity or relative deprivation may impact directly on health, affecting both the endocrine and immune systems. Relatedly, psychosocial stress may cause people to start smoking or to engage in other behaviors which are detrimental to health.

Observational studies, actual experiments, and so-called “natural” experiments have repeatedly shown the power of psychosocial factors on physical health. For example, people in Bristol whose houses were flooded in 1969 were found to have a 50 percent higher mortality rate than members of a control group during the following year.¹⁸ A study of the effects of job losses resulting from a factory closure found that health began to deteriorate when redundancies were first announced—before people actually became unemployed.¹⁹ An Australian study of unemployment found health more closely related to the subjective experience of financial strain than to actual income.²⁰

Workplace studies conducted in Sweden, the United States, Germany, and Britain found that the three most important health-related aspects of the work environment are the amount of control people have over their work, the pressure of work, and the social support they get from colleagues.²¹ These three factors seem likely to apply as much to domestic circumstances as to working conditions. Indeed, numerous studies have confirmed the beneficial ef-

fects on health of more and better social contacts between people.²² Friendships are often seen as protective against the effects of stress. One study showed that they are specifically protective against the stresses of economic hardship.²³ However, poorer people and people in lower status occupations are likely to have less social contact.²⁴

Like social support, most of the psychosocial factors which are deleterious to health are more likely to occur in relatively deprived circumstances. Studies designed to show the links between relative deprivation and health have shown, not unexpectedly, that deprivation causes stress,²⁵ that socioeconomic status affects "fatalism" and attributional style,²⁶ that economic hardship reduces people's ability to fulfill their roles—whether as "breadwinner" or "homemaker"—and can therefore lead to depression.²⁷

It may be that income affects health primarily as a determinant of social status. For example, Eastern Europe (prior to 1989) was an exception to the rule that narrower income differences are beneficial to health. The most likely explanation seems to be that income was often not a good guide to status: in several countries the average manual wage was higher than the average nonmanual wage. It seems that other factors, like party membership, were more important determinants of status.²⁸

It may still seem surprising that income distribution affects us sufficiently to exert a major influence on health through psychosocial pathways. There is, however, no shortage of evidence of the centrality of the sense of social status, feelings of inferiority, and relative deprivation in human social life. Even in early childhood, standards are relative.²⁹ Rather than getting enough care, love, and attention in absolute terms, what is salient in children's minds is any sign that they are less favored than a sibling. If such feelings become part of a pattern they can have a deep emotional impact.

SOME WIDER IMPLICATIONS OF INCOME DISTRIBUTION

If the psychological consequences of wider differences in income are powerful enough to exert a major influence on death rates, it is implausible that they do not also have other social implications. Although no one doubts that poorer areas suffer disproportionately from a wide range of social problems (including drugs, crime,

violence, poor educational performance, and ill health), proving causality is not easy.

Given the controversy over the extent to which the ills of modern society should be blamed on changes in family structure rather than on relative poverty, there is perhaps one international comparison that is worth making. In terms of family structures, Japan and Sweden are at opposite ends of the spectrum. Japan remains closest to the two-parent nuclear family with very few births outside marriage and low divorce rates, while in Sweden over half of births occur outside marriage.³⁰ Yet, these two countries do almost equally well in terms of child welfare and health. What these countries have in common (or had until recently) is a narrow income distribution.

It is also interesting to note that Japan is almost the only developed country which appears not to have been affected by the general rise in crime. In Japan, homicide, robbery, rape, violence, and bodily injury have all shown a long-term decline.³¹ Many Japanese attribute this decline to rapid economic growth, arguing that robbery has decreased because rising incomes mean that fewer people need to steal. This explanation sounds improbable to Westerners who are more likely to believe that improved living standards have somehow made us all more materialistic and selfish. A more plausible explanation is that a component of crime is related less to what has been happening to economic growth and *absolute* living standards than to what has been happening to income distribution and trends in *relative* poverty. While Japan has experienced a long-term decline in income differences, most other developed countries have seen either a worsening or little change in income distribution. Crime in Japan has declined fastest in the large cities, and its association with relative deprivation has weakened. If low *relative* income were a contributing factor, these are the trends which might be expected during a period of declining inequality.

If the relationship between health and income distribution is evidence of a powerful psychosocial effect of relative income, could it be said that each individual's desire for more income is more a desire to improve his relative standing in society than it is a desire for a higher level of material consumption? Such a desire would not be well served by general increases in prosperity, which leave relativities essentially unchanged. Indeed, it would mean that it is not

legitimate to sum individual desires for more income into an aggregate societal demand for economic growth.

Beyond that, however, the psychosocial effect of relative income suggests that the quality of life is better served by reducing relative poverty and narrowing income distribution than it is by further haphazard economic growth. Because health is sensitive to the quantitative as well as qualitative aspects of both material and social change, it could be claimed that it is a better guide to the quality of life than is any mix of economic indicators. Economic indices can only reflect change by reducing qualitative change to quantitative increase, thereby ignoring the social dimension of life. In this context, the fact that in the developed world the predominant position of material factors as determinants of health has given way to social factors has profound implications.

At a time when the environmental impact of economic development is causing increasing concern, it is particularly important to know whether its advantages outweigh its disadvantages. In the developed world the answer to this question can no longer be taken for granted.³² Clearly there are substantial reasons for thinking we need to be more selective about the direction and form which economic growth takes. We continue to think of it as desirable largely because of the short-term discomfort of recession and the need to keep unemployment down and profits up. This is not to suggest that it is time for the industrialized world to bring to a close its two hundred year involvement with economic development. But, rather than seeing innovation as serving the single goal of increasing per capita incomes, we should see it as a means of resolving growing environmental and social problems.

The environment cannot be safeguarded without major technical innovation, just as relative deprivation cannot be diminished except through income growth among the least well-off. Somehow growth has to be put into a new harness. Fortunately, there is increasing evidence that, in addition to the contribution it makes to welfare, income redistribution is also beneficial to productivity and growth. Among OECD countries, those with the fastest increases in productivity tend to be those with narrower income differences.³³ Among seventy rich and poor countries, investment was higher where income differences were narrower. In addition, eight of the rapidly growing East Asian economies narrowed their income differences

between 1960 and 1980,³⁴ and Japan has the narrowest income differentials in the developed world.

This suggests that the pursuit of narrower income differentials may be complementary to growth. Social environments which are less divisive, less undermining of self-confidence, less productive of social antagonism, and which put greater resources into developing skills and abilities, may well turn out to be more innovative and better able to adapt to the environmental problems we face. The health data suggests that the quality of the social fabric, rather than increases in average wealth, may now be the primary determinant of the real subjective quality of human life. With this knowledge let us create high quality social environments where we will be less dependent on the destructive quick fix of indiscriminate economic growth.

ENDNOTES

¹This change was also marked by the decline of mortality from infectious diseases which had been associated with absolute poverty and by the reversal of the social class distribution of heart disease when the so-called "diseases of affluence" became the diseases of the relatively poor in affluent societies.

²Richard G. Wilkinson, "The Impact of Income Inequality on Life Expectancy," in Stephen Platt, H. Thomas, S. Scott, and G. Williams, eds., *Locating Health: Sociological and Historical Explorations* (Aldershot: Avebury, 1993).

³Michael G. Marmot, Abe Adelstein, Nicola Robinson, and Geoffrey Rose, "Changing Social Class Distribution of Heart Disease," *British Medical Journal* II (1978): 1109-12. Steve Wing, "Social Inequalities in the Decline of Coronary Mortality," *American Journal of Public Health* 78 (1988): 1415-16. Aaron Antonovsky, "Social Class and the Major Cardiovascular Diseases," *Journal of Chronic Diseases* 21 (1968): 65-106.

⁴British Medical Association, *Report of the Committee on Nutrition* (London: BMA, 1950).

⁵Ministry of Education, *The Health of the School Child, 1956-7 and 1960-1*, Report of the Chief Medical Officer (London: HMSO, 1958 and 1962).

⁶Pierre Bourdieu, *Distinction: A Social Critique of the Judgement of Taste* (London: Routledge, 1984).

⁷Alison Macfarlane and Miranda Mugford, *Birth Counts: Statistics of Pregnancy and Childbirth* (London: HMSO, 1984). Chris Power, "National Trends in Birthweight and Future Adult Disease," *British Medical Journal* (forthcoming).

⁸John Fox, ed., *Health Inequalities in European Countries* (London: Gower, 1989).

- ⁹Richard G. Wilkinson, "Class Mortality Differentials, Income Distribution and Trends in Poverty 1921-1981," *Journal of Social Policy* 18 (3) (1989): 307-35.
- ¹⁰Richard G. Wilkinson, "Income Distribution and Life Expectancy," *British Medical Journal* (1992): 165-68, 304. Irene Wennemo, "Infant Mortality, Public Policy and Inequality—A Comparison of 18 Industrialised Countries, 1950-85," *Sociology of Health and Illness* 15 (1993): 429-46.
- ¹¹Richard G. Wilkinson, "Income Distribution and Mortality: A 'Natural' Experiment," *Sociology of Health and Illness* 12 (1990): 391-411. Barbara H. Kehrer and Charles M. Wolin, "Impact of Income Maintenance on Low Birthweight," *Journal of Human Resources* XIV (1979): 434-62.
- ¹²Michael G. Marmot and George Davey-Smith, "Why are the Japanese Living Longer?," *British Medical Journal* 299 (1989): 1547-51.
- ¹³Michael G. Marmot and Michael E. McDowall, "Mortality Decline and Widening Social Inequalities," *Lancet* II (1986): 274-76.
- ¹⁴"The rank order of countries in terms of income inequalities strongly corresponds to their rank order in terms of inequalities in mortality. Thus our material corroborates Wilkinson's inferences." Anton E. Kunst and Johan P. Mackenbach, "The Size of Mortality Differences Associated with Educational Level: A Comparison of Nine Industrialized Countries," *American Journal of Public Health* (forthcoming).
- ¹⁵Peter Goldblatt, ed., 1971-81 *Longitudinal Study: Mortality and Social Organisation*, OPCS Series LS 6 (London: HMSO, 1990).
- ¹⁶George Davey-Smith, Martin J. Shipley, and Geoffrey Rose, "Magnitude and Causes of Socioeconomic Differentials in Mortality: Further Evidence from the Whitehall Study," *Journal of Epidemiology and Community Health* 44 (1990): 265-70.
- ¹⁷Department of Social Security, *Households Below Average Income, 1979-1990/1* (London: HMSO, 1993).
- ¹⁸Edward H. Conduit, "If A-B Does Not Predict Heart Disease, Why Bother With It? A Clinician's View," *British Journal of Medical Psychology* 65 (1992): 289-96.
- ¹⁹Norman Beale and Susan Nethercott, "Job-Loss and Family Morbidity: A Study of Factory Closure," *Journal of the Royal College of General Practitioners* 35 (1988): 510-14.
- ²⁰Philip Ullah, "The Association between Income, Financial Strain and Psychological Well-Being among Unemployed Youths," *Journal of Occupational Psychology* 63 (1990): 317-30.
- ²¹Jeffrey V. Johnson and Ellen M. Hall, "Job Strain, Work Place Social Support, and Cardiovascular Disease: A Cross-Sectional Study of a Random Sample of the Swedish Working Population," *American Journal of Public Health* 78 (1988): 1336-42. Robert A. Karasek et al., "Job Characteristics in Relation to the Prevalence of Myocardial Infarction in the U.S.," *American Journal of Public Health* 78 (1988): 910-18. Johannes Siegrist, Richard Peter, Astrid Junge, Peter Cremer, and Dieter Seidel, "Low Status Control, High Effort at Work and Ischaemic Heart Disease: Prospective Evidence from Blue-Collar Men," *Social Science and*

- Medicine 31 (1990): 1127-34. Michael G. Marmot, George Davey-Smith, Stephen Stansfield et al., "Health Inequalities among British Civil Servants: The Whitehall II Study," *Lancet* (1991).
- ²²James S. House, Karl R. Landis, and Debra Umberson, "Social Relationships and Health," *Science* 241 (1988): 540-45.
- ²³Christopher T. Whelan, "The Role of Social Support in Mediating the Psychological Consequences of Economic Stress," *Sociology of Health and Illness* 15 (1993): 86-101.
- ²⁴Peter Townsend, *Poverty in the United Kingdom* (Harmondsworth: Penguin, 1979). Michael G. Marmot, "Social Inequalities in Mortality: The Social Environment," in Richard G. Wilkinson, ed., *Class and Health: Research and Longitudinal Data* (London: Tavistock, 1986).
- ²⁵Whelan, "The Role of Social Support in Mediating the Psychological Consequences of Economic Stress," 86-101.
- ²⁶Blair Wheaton, "The Sociogenesis of Psychological Disorder: An Attributional Theory," *Journal of Health and Social Behaviour* 21 (1980): 100-23.
- ²⁷Catherine E. Ross and Joan Huber, "Hardship and Depression," *Journal of Health and Social Behaviour* 26 (1985): 312-27.
- ²⁸Edmund Wnuk-Lipinski and Raymond Illsley, "Introduction," *Social Science and Medicine* 31 (1990): 833-36.
- ²⁹Judy Dunn and Robert Plomin, *Separate Lives: Why Siblings are so Different* (New York: Basic Books, 1990).
- ³⁰The statistics show that even single parents are well protected from relative poverty in Sweden and their children's health is good. Kathleen Kiernan and Valerie Estauth, *Cohabitation, Extra-Marital Childbearing and Social Policy* (London: Family Policy Studies Centre, 1993).
- ³¹Statistics Bureau, *Japan Statistical Yearbook 1990* (Japan: Management and Coordination Agency, 1990).
- ³²The various attempts to adjust GNPpc to produce better measures of economic welfare have also cast doubt on the benefits of undifferentiated economic growth. In Britain and the United States economic welfare has been declining despite increases in GNPpc. Herman E. Daley and John B. Cobb, *For the Common Good* (London: Green Print, 1990). Tim Jackson and Nic Marks, *U.K. Index of Sustainable Economic Welfare* (Sweden: Stockholm Environment Institute in cooperation with the New Economic Foundation, 1994).
- ³³Andrew Glyn and David Miliband, eds., *Paying for Inequality* (London: Rivers Oram, 1994).
- ³⁴The World Bank, *The East Asian Miracle* (Oxford: Oxford University Press, 1993).

Lurking behind these issues is the abiding problem of the relationship between individual good and common good. American medicine has been based upon the value the society places on the individual; the medical researcher is free to pursue whatever interests and challenges him; the medical patient is free to seek whatever kinds of medical care he thinks he needs, including the freedom to determine what he wants to regard as a medical condition; physicians, patients, and researchers are all free to determine their own value systems and their own ethical codes; everyone is free to get whatever he is willing to pay for, just as everyone is free to live the kind of life he wants, whatever the costs to himself or to society of hazardous living habits. Obviously, a number of practical limitations stand in the way of a full exercise of these freedoms; but there has been, until recently, very little challenge to their abstract validity.

The most important changes which American medicine is likely to witness in the next few decades will come from a systematic questioning of those premises. The beginnings of that challenge can already be seen.

Daniel Callahan

From "Health and Society: Some Ethical Imperatives"
Dædalus 106 (1) (Winter 1977)

The Social Environment and Health

CONSIDER THE ACHIEVEMENTS IN the area of coronary heart disease. For over forty years coronary heart disease has been studied more aggressively than any other disease with the highest level of financial support the world has ever seen. During this period of incredible worldwide effort three risk factors have been identified and universally agreed upon: cigarette smoking, high blood pressure, and high serum cholesterol. Dozens of other risk factors, such as obesity, physical inactivity, diabetes, blood lipid and clotting factors, stress, and various hormone factors, have also been identified but full agreement has not yet been reached. When *all* of these risk factors are considered together, they account for about 40 percent of occurrences of coronary heart disease.¹

How is this possible? Could one or two crucial risk factors have been overlooked? Certainly. But it should be noted that the relative danger of any new risk factors would have to be enormous to account for the remaining 60 percent of occurrences of coronary heart disease. It seems very unlikely that risk factors of such importance would have been simply missed.

Because of this lack of success in identifying risk factors at the individual level, interest in identifying factors in the social environment that may be important determinants of health and well-being is growing. I am not suggesting that the risk factors that *have* been identified are unimportant. On the contrary, they have been very useful in the prevention of coronary heart disease, but, clearly, there are other issues involved that epidemiologists, myself included, do not yet fully understand.

S. Leonard Syme is Professor of Epidemiology at the University of California at Berkeley.

Even when disease risk factors are identified, it is difficult for people to make appropriate behavioral changes. The Multiple Risk Factor Intervention Trial (MRFIT),² an experiment designed to reduce the death rate from coronary heart disease in the United States, is an example of this problem. It was to be the largest, most ambitious, and most expensive experiment ever designed to see if by getting people to change high risk behaviors, the coronary heart disease death rate could be reduced.

Five hundred thousand men in twenty-two cities across the United States were screened for their level of risk for coronary heart disease. Twelve thousand of these men were found to be in the top risk category by reason of their cigarette smoking, high blood pressure, and high serum cholesterol levels. Half of these men worked on lowering their risk in MRFIT clinics and half were sent back to their own doctors. After six years it was found that disappointingly few of the men in the special care group had changed their high risk behaviors.³ To be specific, 42 percent of the smokers stopped smoking, which is probably the best record ever achieved in a smoking cessation program, and about half of the men with hypertension had it under control at the end of the trial. The results in reducing serum cholesterol by dietary intervention are less noteworthy: there was an overall reduction in serum cholesterol levels of 6.7 percent.

While it is evident that some changes took place, the results were disappointing given that those selected were, by all indications, perfect compilers. The limited success that was achieved is, I argue, probably the best we can ever hope for in one-to-one intervention programs. It was the high-water mark. The message from this extreme example is clear and well-known: it is very difficult for people to change high risk behaviors even when they really want to and even when every effort is made to help them.

To make matters worse, there are always new people entering the at-risk population to take the place of those who have changed their high risk behaviors. It is, therefore, clearly evident that despite the efforts of such programs as MRFIT the distribution of disease in the population remains unchanged. This is not surprising considering that nothing is being done to alter the societal forces that caused the problem in the first place.

Epidemiologists tend to look at *individuals* in order to find possible causes of disease even though it is clear that this will not be

helpful in understanding the distribution of disease in the population. This point was forcefully made at the turn of the century by the French sociologist Emile Durkheim.⁴ In his studies of suicide, Durkheim noted that suicide rates exhibited a patterned regularity over time and place. He found that suicide rates are consistently higher in certain countries and in certain groups, even though individuals come and go from those groups. What is puzzling in this is that if the causes of suicide are rooted in the individual, as they are generally thought to be, how can there exist a patterned regularity in groups? Durkheim reasoned that there must be something about the group that promotes a higher or lower rate of suicide in the group as compared with other groups. This factor would not, of course, predict *which* individuals in the group may commit suicide but it would account for the consistency of the *rate*.

This is not to deny the importance of working with individuals, but it is also important to recognize that this approach can have only limited success because it does nothing to alter those forces in society that cause these problems and that continue to provide a fresh supply of at-risk people. The lesson learned from Durkheim is that if epidemiology is to be useful in developing methods of prevention, it will need to study the community and the population.

Almost all epidemiologists study large numbers of individuals in communities. This is not epidemiology. It is clinical medicine in large groups. It may be that the clinical tradition runs so strong that it has come to dominate not only clinical medicine but epidemiology and public health as well. Epidemiologists typically use a *clinical* classification of diseases in their research, studying such diseases as heart disease, cancer, and AIDS. But heart disease, cancer, and AIDS are not public health problems, they are clinical problems. Nevertheless, much of the funding of epidemiological studies and much of the way epidemiologists organize themselves are centered around these clinical disease categories.

Consider the approach that epidemiologists used in the control of infectious diseases. Conditions were divided into such diagnostic categories as waterborne, food-borne, airborne, and vector-borne. This diagnostic classification scheme is of little use in clinical practice since it does not give any insight into treatment possibilities. What this classification scheme *does* is target those aspects of our

environment that are generating disease and suggest at what points in the environment interventions might be deployed.

In short, epidemiologists have adopted a disease classification scheme from the clinical model that focuses attention on clinical states of health and that draws attention to the affected individual. When social, community, and environmental forces are considered—and, of course, they are—they are thought to shed light on the individual and his health. The focus, however, is not where it ought to be.

There are at least two reasons why it is so difficult to think in terms of the social environment. As indicated earlier, the clinical tradition in epidemiology is so pervasive that it overwhelms all other approaches. For example, one can apply for money from the National Institutes of Health (NIH) to study arthritis, metabolic disease, heart disease, cancer, and eye disease, but one cannot request funds to study health. Only the study of diseases of clinical relevance will be supported. While this categorical support concerns me, I realize that not having a well-reasoned and useful set of categories to offer hinders our progress.

The second reason that it is difficult to think about the social environment as a determinant of health is more subtle but perhaps even more important. It may be that a community approach to infectious disease is easier to develop because most of us are exposed to infectious agents whether or not we want to be. A contaminated water or food supply puts all of us at risk, as does toxic air or infected mosquitoes. These clearly are public health issues and should concern public health agencies. In contrast, the way we eat, drink, smoke, drive, sit, run, and work are private behaviors and as such are not in the jurisdiction of public health. In this view, each of us is responsible for the heart disease, cancer, and AIDS that we may develop.

As individuals, we *are* ultimately responsible for our own health and our own behavior. But it is naive to think that we are free agents in this. All of us are influenced by forces in the community that shape our choices and preferences. For example, in a project I am involved in that is being conducted in Richmond, California to encourage the cessation of cigarette smoking, we have been struck by the enormous pressures put on teenagers to smoke. The messages they receive from advertising and their peers vividly portray smok-

ing as positive and glamorous. As a result, many teenagers report that it is far easier to smoke than it is to refrain from smoking. Similar pressures exist for violent behavior, the use of alcohol and drugs, and casual sexual behavior. And the influence of such community forces is not limited to children and teens. In a study I am involved in with San Francisco bus drivers, dramatic differences have been observed between bus barns in the prevalence of alcohol abuse, obesity, cigarette smoking, and absenteeism. These differences are shaped more by community pressures in each barn than by individual differences between drivers.

Just as bad water and food may be harmful to our health, unhealthful forces in our society may be detrimental to our capacity to make choices and to form opinions. It makes no sense to ignore a contaminated water supply just as it makes no sense to urge people to change their unhealthful behaviors when we leave untouched those forces that prompted much of that behavior in the first place.

Following the work of Durkheim, it may be beneficial to take advantage of patterned regularities, particularly those of social class. We have known for hundreds of years that people in the lowest levels of the social class hierarchy have the highest rates of virtually every disease and condition.⁵ Despite universal recognition, epidemiologists have not studied the reasons for this phenomenon. The list of possible explanations is long: poverty, substandard housing, unemployment, poor nutrition, inadequate medical care, and limited education, to name only a few. Instead of studying social class epidemiologists typically "hold it constant" so that other things can be examined. If we did not, social class, with all its various components, would obscure the roles of all other factors. Consequently, we know little about the various subcomponents associated with social class.

But there is another, even more important, reason for our unwillingness to study social class: we do not feel that anything can be done about it. Social class is a product of vast historical, economic, and cultural forces, and, short of revolution, it is not something one targets for intervention. So we give up and instead urge people to lower their fat intake. But this view is not based on fact. For example, if research were to show that people in the lower social classes had higher rates of disease because they were poor, it might

be true that interventions would be difficult. In reality we have no evidence that lack of money is the major culprit. It seems premature to conclude that social class is too difficult to consider or deal with.

An interesting example of how we might better approach this issue is provided in the work by Marmot and his group on British civil servants.⁶ They show that those at the very bottom of the civil service hierarchy have heart disease rates four times higher than those at the top. After adjusting for such risk factors as blood pressure, serum cholesterol, smoking, and physical activity, the difference between these groups is still threefold. They also show that those one step down from the top of the hierarchy, civil servants who are professionals and executives, such as doctors and lawyers, have heart disease rates that are twice as high as those at the very top, the upper-class directors of agencies.

It is not surprising that those at the bottom would have higher rates of disease than those at the top. What is surprising is that doctors and lawyers, only one step removed from the top, also have higher rates. Marmot and his group found that there is a gradient of the rate of disease from the top of the British civil servant hierarchy to the bottom.

One might be inclined to think that this phenomenon is somehow unique to the British civil service. It is not. A similar gradient can be found almost everywhere in the world for virtually every disease that has been studied.⁷ The problem posed by this finding is this: We can imagine why those at the bottom have higher rates of disease but how can we explain a gradient? How is it that those one or two steps from the top have higher rates of disease than those at the top even though they do not suffer from the problems experienced by those at the bottom?

Perhaps it is because as one moves down the social class hierarchy, one has increasingly less control of one's destiny, less opportunity to influence the events that affect one's life.⁸ This hypothesis is general and unproved, and it does not specify whether control involves money, power, information, prestige, experience, or something else. Over the years many social scientists have studied other concepts related to the idea of the "control of destiny," including mastery, self-efficacy, locus of control, learned helplessness, ability to control, predictability, desire for control, sense of control, powerlessness, hardiness, and competence. It may be of value to look

for common denominators in that body of work. It has been shown repeatedly that the more effectively people deal with forces that affect life and living circumstances, the better their health and well-being. If this idea, or something like it, were supported by research, it might open an avenue for intervention.

In support of this notion, a group of epidemiologists led by Robert Karasek⁹ from the United States and Tores Theorell¹⁰ from Sweden has shown that rates of coronary heart disease are higher among workers who are in high demand jobs and who have limited discretion and latitude for dealing with those demands. The work of these researchers is especially impressive because previous studies of job stress failed to establish a link between job pressures and health even though this issue had been examined intensively.

These findings may relate to our study of San Francisco bus drivers by illustrating the usefulness of an environmental perspective in studying health and disease and the possible value of the concept of the "control of destiny" in developing interventions to prevent disease. Several previous studies have noted that bus drivers generally have a higher prevalence of hypertension, as well as diseases of the gastrointestinal tract, respiratory system, and the musculoskeletal system, relative to workers in other occupations.¹¹ Based on these findings, it has been suggested that certain aspects of driving a bus may create an increased risk of disease for workers in that occupation.

From a clinical standpoint it is valuable to identify individual drivers who have diseases so that they may be treated. It would also be valuable to teach drivers about better posture, more healthful eating habits, and alternative ways of dealing with job stress. However, from an environmental perspective, it would perhaps be more useful to identify those aspects of the job that might be changed to prevent disease.

In conclusion, there seems now to be enough evidence to warrant a serious research program on the social environment as an important determinant of health. Our work with individuals has been useful and productive, but this approach alone clearly will not lead to an effective program of health promotion and disease prevention. A new initiative focusing on the environments in which we live must now become a priority for us all.

ENDNOTES

- ¹Michael Marmot and Warren Winkelstein, "Epidemiological Observations on Intervention Trials for Prevention of Coronary Heart Disease," *American Journal of Epidemiology* 101 (1975): 177–81.
- ²Multiple Risk Factor Intervention Trial Research Group, "The Multiple Risk Factor Intervention Trial," *Preventative Medicine* 10 (1981): 387–553.
- ³Multiple Risk Factor Intervention Trial Research Group, "The Multiple Risk Factor Intervention Trial—Risk Factor Changes and Mortality Results," *Journal of the American Medical Association* 248 (1982): 1465–476.
- ⁴Emile Durkheim, *Suicide: A Study of Sociology*, ed. and trans. George Simpson (Glencoe, Ill.: Free Press, 1951).
- ⁵Aaron Antonovsky, "Social Class, Life Expectancy and Overall Mortality," *Milbank Memorial Fund Quarterly* 45 (1967): 31–73; S. Leonard Syme and Lisa F. Berkman, "Social Class, Susceptibility and Sickness," *American Journal of Epidemiology* 104 (1976): 1–8; and Evelyn M. Kitagawa and Philip M. Hauser, *Differential Mortality in the United States* (Cambridge, Mass.: Harvard University Press, 1973).
- ⁶Michael G. Marmot, Geoffrey Rose, Martin Shipley, and P. J. S. Hamilton, "Employment Grade and Coronary Heart Disease in British Civil Servants," *Journal of Epidemiology and Community Health* 3 (1978): 244–49.
- ⁷Mary N. Haan, George A. Kaplan, and S. Leonard Syme, "Socioeconomic Status and Health: Old Observations and New Thoughts," in John P. Bunker, Deanna S. Gomby, and Barbara H. Kehrler, eds., *Pathways to Health: The Role of Social Factors* (Menlo Park, Calif.: Henry J. Kaiser Family Foundation, 1989).
- ⁸S. Leonard Syme, "Control and Health: A Personal Perspective," in Andrew Steptoe and Ad Appels, eds., *Stress, Personal Control and Health* (New York: John Wiley, 1989).
- ⁹Robert Karasek, Dean Baker, Frank Marxer, Anders Ahlbom, and Tores Theorell, "Job Decision Latitude, Job Demands, and Cardiovascular Disease: A Prospective Study of Swedish Men," *American Journal of Public Health* 71 (1981): 694–705.
- ¹⁰Tores Theorell, Lars Alfreddson, Sarah Knox, Aleksander Persk, Lennart Svensson, and Dick Waller, "On the Interplay between Socioeconomic Factors, Personality and Work Environment in the Pathogenesis of Cardiovascular Disease," *Scandinavian Journal of Work, Environment and Health* 10 (1984): 373–80.
- ¹¹Marilyn A. Winkelby, David R. Ragland, June M. Fisher, and S. Leonard Syme, "Excess Risk of Sickness and Disease in Bus Drivers: A Review and Synthesis of Epidemiologic Studies," *International Journal of Epidemiology* 17 (1988): 124–34.

Monitoring Motherhood: Sociocultural and Historical Aspects of Maternal and Child Health in Japan

INTRODUCTION

STATISTICS ABOUT INFANT MORTALITY, namely the Infant Mortality Rate (IMR), the number of deaths within one year after birth (per thousand births), are usually taken to be one of the most reliable indicators of health in any given society. Consequently, public health activities in most parts of the world have targeted the prevention of infant deaths, and researchers have put serious efforts into assessing how this goal can be achieved. Recent investigations have demonstrated the significance of sociocultural factors in the reduction of IMR.¹ Perhaps the most significant factor is the wealth of the nation. In principle, the higher the per capita gross national product (GNP), the lower the IMR; the IMR of so-called developing countries is estimated at more than ten times that of developed countries. A recent study on significant infant mortality reductions in several regions showed the importance of the following factors: female autonomy, education (especially female), access to health-care services and mechanisms to guarantee their efficient operation, an adequate minimal standard of nutrition, universal immunization, and a commitment to the establishment of effective perinatal services.² Other studies in North America have shown that the distribution of wealth across social class, access to prenatal care,

Naoko T. Miyaji is a practicing physician and lecturer in the Department of Hygiene, Kinki University Medical School, Osaka-sayama city, Japan.

Margaret Lock is Professor of Medical Anthropology in the Departments of Social Studies of Medicine and Anthropology at McGill University, Canada.

teenage pregnancy, out-of-wedlock birth, the presence of minorities who experience discrimination, immigrant and refugee status, and low educational level of parents also influence the IMR.³

For those researchers interested in health demographics, Japan has attracted attention. Since 1984, the Japanese IMR has been the lowest in the world.⁴ In 1992, the IMR was 4.5 (as opposed to 6.8 in Canada in 1991 and 9.2 in the United States in 1993), and the Perinatal Mortality Rate was 5.2 (3.5 late fetal death rate plus 1.8 early neonatal death rate) (as opposed to 8.0 in Canada in 1989 and 10.1 in the United States in 1988). Japan also boasts the longest life expectancy at birth: 76.25 years for men and 82.51 years for women (1993).

PREGNANCY AND CHILDBIRTH IN JAPAN: CURRENT ATTITUDES AND PRACTICES

Reproductive Histories and Cultural Expectations

Demographic figures in Japan show a strikingly uniform pattern of reproduction. Almost all children are born after marriage; the majority of women stop working once pregnant for the first time; and most women have had two children, all they plan to have, by the age of thirty-five. This trend indicates the existence of shared expectations among the public about the correct conditions for childbirth and the raising of children, and, at the same time, makes it easier for the government to manage maternal and child care.

The average age of marriage is 26 for women and 28.4 for men. The divorce rate is increasing, but remains low at 1.45 per thousand, much lower than in America (4.7 per thousand). The average interval between marriage and the birth of the first child is two years. Ninety-nine percent of women who gave birth in 1992 were married, a figure which has been stable for three decades. In Japan, 89.9 percent of babies are born to mothers between the ages of twenty and thirty-five. The numbers of births by teenage mothers have been increasing slightly but remain low at 1.5 percent of all births. Delivery by women aged thirty-five to forty is 7.6 percent of all births, and 1 percent by women aged forty or older. The age distribution for giving birth has not changed significantly for over a decade. A well recognized risk factor for infant mortality is low

birth weight. The rate of low birth weight babies (less than 2,500 grams) is 6.2 percent for boys and 7.5 percent for girls, not a particularly impressive statistic (in the United States the rate was 7 percent in 1990).

Japanese women have a low fertility rate: the birth rate is 9.8 per thousand and the total fertility rate (the average number of children a woman is expected to produce during her reproductive period) is 1.5 (in the United States it is 2.01). Only 3.1 percent of Japanese babies are fourth or lower in the birth order. The three major causes of infant deaths are congenital anomalies (37.2 percent), birth trauma (17.9 percent), and accidents (6 percent). Overall, neonatal deaths comprise 53 percent of all infant deaths in Japan.

Monitoring and Preventive Care

In Japan pregnancy is carefully monitored both as part of the medical system and in the privacy of the home, and there is a comprehensive system of preventive care. Public Health Centers (*hokensho*) organize most preventive services in cooperation with hospitals and clinics, and almost all women comply with this monitoring system. Monitoring starts with the registration of the pregnancy with the local government, which is required by the Maternal and Child Health Act. More than 90 percent of pregnancy registration is done before the twentieth week of pregnancy,⁵ which means that most women learn of the pregnancy soon after conception, visit a doctor, and then comply with professional supervision while receiving support from public health services.

One important device for careful monitoring is the Maternal and Child Health Handbook (*Boshitecho*), which is handed out after registration of the pregnancy. This small book becomes the child's official health record from the prenatal period until he or she enters primary school at age six. Pregnancy and delivery status, the developmental curve of the child, vaccinations, and results of health checkups are recorded and referred to at follow-up examinations. One other purpose of the Handbook is to provide information to women about normal pregnancy, dangerous symptoms during pregnancy, developmental stages of children, and information on clinics and available services at public health centers near their places of residence. The Handbook also has symbolic meaning, stimulating an awareness in the woman that she will shortly become a mother.

Women report that they feel both joy and anxiety and become sensitive about future maternal responsibilities once they receive the Handbook.⁶ The Handbook is often passed on to the adult child, becoming a valued childhood memento.

Much information is also available from commercial guidebooks and magazines. Currently, two monthly magazines deal with pregnancy and childbirth, one of which sells about 270,000 copies every month. The number of magazines about parenting are more numerous, and each one specializes in children of a certain age. Magazines encourage a positive attitude toward pregnant life and are packed with visual images. Although medical authorities often criticize these magazines as biased and scientifically unsound, they cover topics of interest to women which are rarely dealt with by professionals: unusual methods for delivery, prior experiences of other women, sex during pregnancy, maintenance of weight, education about fetal development, financial plans for delivery, introduction of popular obstetrical clinics, and so on. These magazines also provide medical information about such things as hepatitis B and rubella, delivery in mid-life, medication during pregnancy, and explication of medical terms.⁷

A prenatal checkup is recommended monthly until twenty-seven weeks of pregnancy, biweekly to thirty-five weeks, and weekly after that. Ultrasound is used routinely to monitor the development of the fetus. Other invasive tests such as amniocentesis are rarely used. Before and after birth, public health center activities include regular checkups, vaccinations, home visits, and group guidance. Public health nurses are the major resource for these activities. Maternity classes (*hahaoka kyōshitsu*) and child-care classes (*ikuji kyōshitsu*) are designed not only to provide practical knowledge but also to facilitate mutual support among mothers. Home visits are available for high-risk pregnancies and for families with neonates and low birth weight babies.

Almost 100 percent of deliveries are carried out in medical institutions including hospitals, clinics, and maternity homes, and 98 percent are under a doctor's supervision. Delivery rooms usually have a full range of medical technology on hand: a fetal monitor is used during birth in the majority of cases, ostensibly to reduce fetal death and malformation. Making individual choices in connection with birth is attracting the attention of increasing numbers of

mothers, but still there is little variation in delivery methods; neither childbirth with full anesthesia nor "natural" childbirth with only the assistance of a midwife is common. The rate of cesarean sections (10 to 20 percent of all births depending on the institution) is significantly lower than in North America. Hospital stay after delivery is usually one week—longer than in North America. New mothers usually stay in the hospital in rooms apart from their babies and midwives are on hand to give advice. Since 1975, breast-feeding has been promoted nationwide; 43 percent of mothers breast-feed while another 43 percent use both breast milk and artificial milk. Since 1958, low birth weight babies are registered, and the cost of inpatient care is publicly financed. Neonatal Intensive Care Units and Perinatal Intensive Care Units are distributed throughout the country. A comprehensive vaccination program exists, and preventive measures against the transmission of hepatitis B have been implemented since 1985. A system of screening was also developed for treatable congenital diseases, including phenylketonuria, maple syrup disease, homocystinuria, hystizinemia, galactosemia, hypothyroid disease, hyperadrenal disease, and neuroblastoma. The start of this program was later than in many other countries, but the coverage is virtually universal. Participation is, in theory, voluntary, but it is assumed that no one will refuse, and since 1984 the participation rate has been over 99 percent. Besides frequent health and developmental checks in clinics during the first year, special development checks are required by law for all children at eighteen months of age (since 1977), and again when the child is three years old (since 1961).

Once a child is enrolled in school, the school health-care system provides comprehensive health education, an exercise program, nutritional management through a school lunch program, yearly checkups, and other care. There are also public services, including clinics and telephone counseling services, for adolescents and young adults and their parents which are designed to raise the consciousness of young people about the importance of correct reproductive practices.

Public Support for Birth and Infant Health Care

Researchers outside Japan are often perplexed by the combination of excellent health indicators and relatively low health-care costs in

Japan. This is partly explained by the fact that normal pregnancy and delivery are excluded from the National Medical Expenditures. Maternal and child health (MCH) activities are covered under various laws by several ministries and at different levels of organization (national, prefectural, city). Thus, the extent of expenses for MCH is not easily established.

Japan has an obligatory universal medical insurance system. Although expenses for normal delivery (around \$30,000-\$34,000, calculated at \$1/¥100) are not covered by this system, there is a system of reimbursement. A monthly allowance is given to a family with a child under three years of age (\$50 for the first and second child, \$100 for the third child and subsequent children). In addition to the coverage for group support and home visits by *hokensho*, two health checks during pregnancy and further checks in connection with the child's development are publicly supported. Inpatient care for low birth weight babies and treatment of children with disabilities and chronic diseases are also publicly covered. Financial assistance is available for the treatment of pregnant women with gestosis and other medical complications. If a mother is economically disadvantaged, pregnancy and delivery are financially covered by the Child Welfare Act. A third act, the Livelihood Protection Act, also pays medical fees for low income households. This assistance minimizes the influence of poverty on the health of mothers and children. However, because wealth is distributed relatively equally nationwide, few families find themselves in great need of assistance.

MCH in Everyday Life

In Japan, pregnancy and childbirth are considered family events and not only experiences for the young couple. The new baby is the successor to the extended family heritage and bloodline. Parents (especially the mothers) of both new parents are deeply involved in the process of pregnancy and early socialization, and pregnancy is understood as a rather precarious and special time by all family members.

A woman's sense of responsibility is facilitated by various traditional customs, usually initiated by the extended family. *Omiyamairi* is the custom of going to pray at a Shinto shrine which "specializes" in safe childbirth, where a special *obi*, a pregnancy sash, is pur-

chased and wrapped around the abdomen of a pregnant woman. This sash is believed to protect the uterus from the cold, keep the fetus stable, and prevent it from becoming too big.⁸ Through these rituals a maternal identity is established. At the same time both material and nonmaterial resources, such as emotional support, are mobilized for the mother and baby by family members.

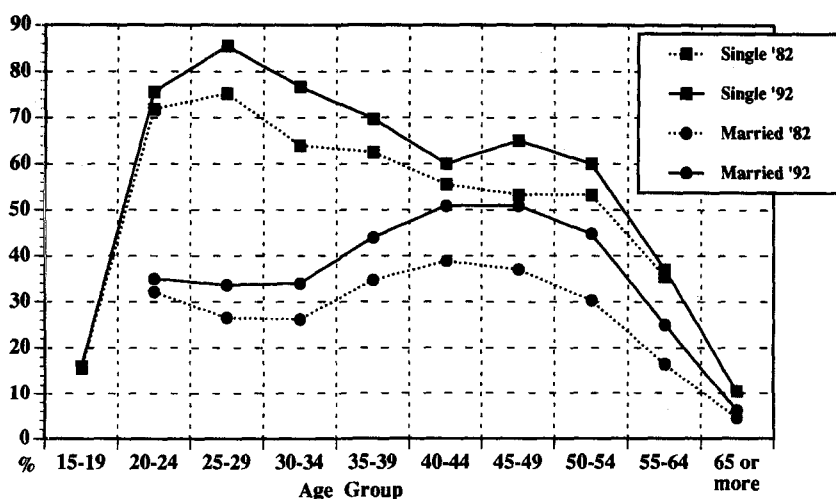
Mothers and Work

The work environment of pregnant women is considered to be very important, and regulations about maternal leave (fourteen weeks) and prohibition of dangerous and harmful work come under the Labor Standard Law.⁹ In 1992, a law about child-care leave was enacted. As a result, a leave of up to one year is allowed for both women and men, though men very rarely take a leave. Usually child-care leaves go unpaid or are given with only very limited pay, but social security and health insurance are covered by about half of the major Japanese enterprises. In 1992, 34.5 percent of mothers took child-care leave, and 61 percent of them took up to six months.¹⁰

The occupational health of mothers has not been seriously considered until recently because the majority of women were expected to stop working once they became pregnant with their first child. As shown in *Figure 1*, the employment rate of women, particularly married women between the ages of twenty-five and thirty-five, is low. The majority of women give priority to their "duty" to produce and raise a healthy child. This may be the result of social pressure or an internalized ideal about "correct" behavior. According to a 1991 survey, 31.2 percent of women leave their jobs once they are pregnant. A survey conducted in 1992 showed that only 26.3 percent of women and 19.8 percent of men think it is better for women to continue to work once pregnant; 2.8 percent of women and 5.7 percent of men think women should not work at all; 21.9 percent of women and 29.9 percent of men think women should quit working for good once they get married or become pregnant; and 45.4 percent of women and 39.2 percent of men think women should stop working at least until their children reach primary school age. Enrollment for nursery care is low for the first three years: 2 percent of infants under the age of one, 9 percent of one year olds, 15.5 percent of two year olds, and 51 percent of three

year olds. Enrollment jumps to 90.3 percent for four year olds because by this age children begin to receive education which will prepare them for the competitive Japanese school system. Because child-care facilities for younger children are very limited, working women often rely on their parents for assistance. Ministry of Labor surveys show that 67.2 percent of working women who have children under one year of age turn to grandparents for child care, and 28.3 percent opt for nursery care.

FIGURE 1. Rate of Working Women by Age Group



Although the younger generation tends to be positive about female employment, the shared value that prohibits mothers with small children from working remains strong. Not only pregnancy and childbirth, but also child care and all other housework are still usually considered to fall entirely into the woman's domain. In the survey mentioned above, 55.6 percent of women and 65.7 percent of men agreed that "men's priority is work and women's priority is the family." Eighty-three and nine-tenths percent of women and 87.7 percent of men agreed that "it is acceptable for women to work but they should be responsible for the housework and child care." Another survey showed that whereas women spend on average 3 hours and 52 minutes on housework, men spend on average only 24 minutes per day, and the time is even shorter (19 minutes)

for men whose wives work outside the home. *Satogaeri shussan*, a custom that pregnant women return to their natal homes for child-birth, not only indicates strong family ties but also a practical need because the help of a husband cannot be relied on. Today, however, more women stay in their own homes and are looked after by their mothers or mothers-in-law.

The Socioeconomic Situation

Let us now turn to more general socioeconomic conditions. Factors such as a nation's wealth, high educational levels, a low rate of unemployment, a relatively equitable distribution of wealth, and little ethnic conflict facilitate a generous investment into maternal and child health on the part of both the government and the family and make these public health activities more effective.

The gross national product per capita (GNPpc) in Japan has increased for four decades and is now the second highest in the world. The unemployment rate has been stable at around 1 to 3 percent for four decades. In Japan the lowest earning fifth of all households has a higher share of total household income than comparable households in all other countries reporting to the World Bank (8.7 percent in 1979, 10 percent in 1989); the highest earning fifth of Japanese households has the third lowest share, comparatively speaking, of total household income (37.5 percent in 1979, 34 percent in 1989).¹¹

Compared to North America, where ethnic minorities, immigrants, and refugees make up a large percentage of the population, ethnic uniformity remains high in Japan. Although the number of foreigners is increasing, the percentage of births among non-Japanese couples is 0.69 percent of all births, and births among Japanese married to non-Japanese is 1.5 percent. This goes far to explain why it is possible to have a widely shared value system in connection with reproduction.

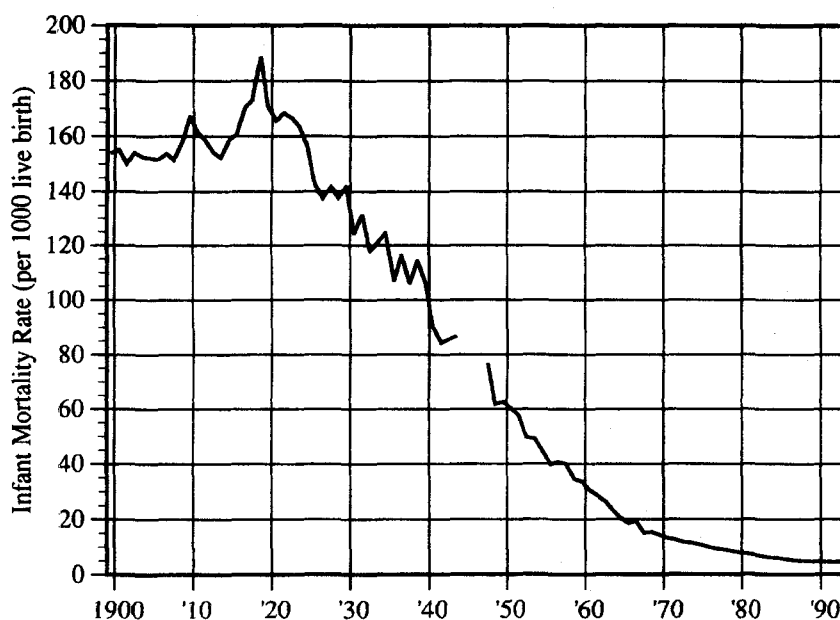
THE ESTABLISHMENT OF MATERNAL AND CHILD HEALTH AS A GOVERNMENT CONCERN

It is generally accepted that improvement in the health of the Japanese population started after World War II and was made possible by advances in medical knowledge and practice. However,

health indicators, including the IMR, had started to improve much earlier. An analysis by Nishida showed that the contribution of medical technology to the decline of the death rate between 1920 and 1980 was less than 30 percent for infectious diseases, and at most 20 percent for infant-specific noninfectious diseases.¹² This study replicates Thomas McKeown's well-known analysis which showed the greater importance of socioeconomic changes for improvement in health.

Figure 2 shows that the IMR was about 160 per thousand births in the Meiji period (1868–1912), and that it decreased to under 100 in 1940 and continued to drop despite the war. In the postwar years the IMR dropped from 76.7 in 1947 to 10 in 1975, and became the lowest in the world in 1984.

FIGURE 2. Trends of IMR in Japan (1899–1992)



Improvements in general health status in Japan occurred concurrently with other social changes. In the past one hundred years Japan has experienced extremely rapid economic growth. At the same time, the political structure, family life, educational system, and status of women have changed significantly, particularly fol-

lowing the Meiji Restoration and again after defeat in World War II. Public health policies and hygiene and nutritional practices were all modified as well. To what extent each factor has contributed to the actual improvement of health is difficult to estimate.

Hygiene campaigns, provision of universal education, and support of the patriarchal family system were central to the modernization policy of the Meiji government. Later, a protective policy for MCH was initiated to produce a strong military force early in the Showa period (1926–1988). Drastic reforms were carried out at this time, based on innovative ideas about the importance of a healthy population. Later the foundation for the current health system was laid down by the American Occupational Force in the postwar era.

Although each of these policies was created for a specific purpose and was colored with the dominant ideologies of the time, they also demonstrated a long-term vision which emphasized human development and well-being. Most importantly, the implementation of policies was systematically carried out from the beginning of the Meiji era by means of campaigns to reform the “habits” of the people through “moral suasion” (*kyôka*).¹³ The “masses” were educated to “rationalize” their lives, and women’s groups and public health workers ensured that women in their locality cooperated in the name of progress, a situation which continues today, although the nationalistic underpinnings are no longer visible.

The Meiji Restoration

The Meiji Restoration of 1868 ensued upon a 250 year closure of Japan to the outside world. The new government was anxious to bring Japanese standards up to the level of the West through increased industrialization and an imitation of what was understood as a “western” way of life. The ideas of democracy and of Enlightenment philosophy were advocated by the Meiji leaders. The occupational caste system of the Edo period was abolished, and all citizens were recognized for the first time as equals. The writings of John Stuart Mill and Rousseau and the French Declaration of Human Rights were introduced, but for ordinary people in the villages the reform of daily life had a more immediate impact. Both the central and local governments made detailed regulations about community practices, including the prohibition of costly wedding

ceremonies, marriages at an early age, easy divorce, and remarriage. Some reforms were designed to liberate women.¹⁴ For example, the government abolished the prohibition of entry of women to areas designated as sacred and allowed them to pass through certain travel checkpoints for the first time.

From the outset an explicitly scientific approach to health was advocated, including an emphasis on improved hygiene, and preliminary experiments with vaccination were carried out. Abortion by nonlicensed birth attendants (*sanba*) and sales of abortion-inducing drugs were banned, and registration of pregnancies was started. This regulation was designed to abolish the practice of *mabiki* (infanticide—literally “thinning-out”), resorted to particularly in times of famine or when an infant was born with an obvious defect.¹⁵ Hygiene campaigns were organized by voluntary community groups and the importance of women’s participation in hygiene improvement was emphasized in magazines. Women’s groups were organized to incite a change of practices in connection with bathing, dress, menstruation, child raising, water use, infant safety, and the education of children at home. By 1913, there were more than two thousand such groups in Japan.

Education

The Meiji government was conscious that “human development,” that is, improvement in the health and abilities of citizens, was important in order to compete with imperialist countries in the West. Investment in education at this time provided an essential base for implementing subsequent changes. If, for example, the general educational level had not been improved in the Meiji period, the Taisho democracy, the women’s movement, and economic recovery after the war may have been hampered. And, in all probability, public health activities would not have been so effective.

The first state school system, educating both men and women, was established in 1873. At first more than half of eligible boys but only about one-fifth of girls were enrolled in primary school. A change in the school ordinance in 1900 made tuition essentially free, and a stronger obligation was placed on parents and employers to guarantee school attendance for children. Female education was emphasized because it was understood that “women are the educators of children.” Primary school enrollment was 90 percent

for boys and 72 percent for girls in 1900, and by the end of the Meiji era (1912) it had reached 97.6 percent for girls and 98.8 percent for boys. In all schools, hygienic practices and moral education were considered central to the production of mature adults who would make a contribution to the State. The Emperor was symbolic of national unity, not political and legal unity, but patriotic and civic. Loyalty to the Emperor, and by implication the State, was associated with progress, an ideology which was taught in school.¹⁶

The actual enrollment rate in schools varied according to region and social class. However, the idea that all children needed an education became widely accepted at this time. Higher education, too, became popular following mass migrations from villages to cities during the early part of this century.

The Meiji Family System and the Position of Women

Prior to the Meiji period, family relationships had been remarkably flexible in Japan. In seeking to create a centralized state, the Meiji government instituted the formal extended family (*ie*) system which until then had only been in use among the samurai class (4 percent of the population). In line with the Civil Code promulgated in 1899, all marriages were formally arranged by the extended family. Once married, a woman left her natal family and was incorporated into that of her husband. A woman's most important duty was to produce a son who would be the family successor, and if she did not do so within three years of marriage, divorce was usual. Although the code required monogamy, the custom of men to keep mistresses was socially acceptable. Extramarital affairs were considered adultery only in the case of wives, and women could not initiate divorce. In fact, without consent of a father or husband, women could not take legal action of any kind.

The Emperor was designated by the government as the pinnacle of the formal family system, thus making obedience to the government a moral issue. Order and stability were secured primarily by an authoritarian regime established in the home, and, because the head of the family had a duty to support family members, the government felt little obligation to establish a welfare system. This patriarchal ideology was promoted through the formal education system. The moral teaching known as "*ryôsai kenbo*" (good wife

and wise mother) emphasized the importance of chastity, obedience, and patience, as well as a woman's devotion to her husband, his parents, and her children.

Counter to the official code, women's rights were explicitly advocated from the early Meiji period, and arguments against the authoritarian family system were widely voiced. Certain women became active in the Women's Movement, and the focal point of the heated debate in 1918 about the "protection of motherhood" (*bosei hogo ronsō*) was the relationship of women's economic independence to their maternal "duty," a theme still relevant today. A movement toward women's suffrage was also active at this time. But political oppression by the government intensified with militaristic expansion, thus keeping women firmly in a position of subordination.

Nationalism and Imperialism

In the Meiji period, virtually no preventive measures were taken in connection with MCH, with the exception of several hygiene campaigns. In Taisho (1912–1926), however, the systematic collection of vital statistics was first carried out, including surveys which revealed the dismal health of the Japanese by international standards.¹⁷ Despite the existence of a widespread belief that prevention of infant mortality was contradictory to natural selection and preventive measures would simply increase the presence of weak people in society, public health officials began to advocate the reduction of the IMR and started relevant epidemiological investigations. By 1926, the government had recommended the establishment of child health-care centers, which was partly realized when the Public Health Center Act (*Hokensho Ho*) was enacted in 1937. Relevant research and practices in rural areas, a traveling midwifery system, and pediatric clinics (*jidō sōdansho* and *nyūyōji sōdansho*) were institutionalized at this time. In addition, educational activities focusing on the impact of poverty on health were organized by socialist groups. Responding to the grave condition of workers' health in factories, the Factory Act (*Kōjō Hō*) was passed in 1911 and enacted in 1916, although the actual application of the law was not very effective for many years. This act regulated working hours for women and young people in industry and prohibited the employment of children under the age of ten.

As Japan entered into an imperialist and expansionist era, the creation of a strong defense force became very important to the government. The army's survey of draft-age men in 1936 revealed that a high percentage of men did not qualify for military service because of malnutrition, infectious diseases, or job-related disabilities. Their poor physical condition in comparison to their Western counterparts motivated the government to pay attention to the health of its citizens. The government explicitly stated that children are the source of the future success of a nation and that women should devote their energies to producing many healthy children. Accordingly, awards were given to those couples who raised more than ten healthy children. In 1937, the Mother and Child Protection Law (*Boshi Hogo Hō*) was formalized, and the training of birth attendants (*sanba*) was facilitated during the 1930s to improve the number of healthy births. The Ministry of Health and Welfare was created in 1938 to focus on health and population issues. That same year, medical insurance coverage was expanded. In 1941, a National Eugenics Law passed which prohibited abortion, and, in 1942, the Maternal Handbook was created. Protection of pregnancy was emphasized by the provision of rations for pregnant women.

As the war dragged on and military expenses increased, financial resources for the promotion of health and hygiene diminished. Military expenses consumed more than half of the total national budget in 1935, and more than 80 percent in 1938.¹⁸ To keep the economy from crumbling, women were mobilized as part of the work force, in addition to managing the family.

The start of a serious MCH policy in Japan was part of a national interest in progress (especially economic development) and in the continuity of the extended family system with its associated moral foundation. Women were easily subjugated by the government for these purposes because most women were eager to cooperate with policies which they believed contributed to their own health and that of their families.

Postwar Reforms

After Japan's defeat in World War II, the Army of Occupation, almost entirely American, came into power. Its aim was to demolish the social structures which had nurtured militarism in Japan and to

democratize Japanese society. America intended to reconstruct Japan as an ally to act as a political buffer in Asia. The reforms were all encompassing: land reform enabled more than 80 percent of farmers to own their land, the wealthy nepotistic business conglomerates (*zaibatsu*) were destroyed, and labor laws were introduced.

Various health programs were implemented to deal as rapidly as possible with starvation and the many diseases associated with poverty. Control of infectious and sexually transmitted diseases was given priority. School lunch programs were instituted, and supplementary rations were again supplied to pregnant women. Other programs were designed to produce long-term benefits: a centralized system for the collection and monitoring of health statistics, the retraining of health professionals, and educational reform in schools of medicine and nursing. Changes in nutritional habits were also strongly encouraged. Public Health Centers were reestablished all over Japan as the central basis for primary care. The Child Welfare Law was enacted in 1947, and the Ministry of Health and Welfare created an MCH division at that time to deal with the many war orphans.

As a result of the shortage of resources after the war, rapid population growth became a serious concern for the government. Abortion was made legal in 1948. Under the Eugenics Protection Law, women could abort if the fetus was physically deformed, for economic reasons, or if continuance of pregnancy or childbirth might cause serious harm to the health of the mother. This law is still in effect today, with the result that almost any request for an abortion can be met. Between 1948 and 1955, more than a million abortions were performed each year,¹⁹ the highest rate ever. At that time, a vigorous family planning campaign was introduced, but people were not educated about new methods of contraception.

The majority of health-care reforms implemented by the Occupation forces remain in the current health-care system. Although certain actions ordered by the Occupation caused resentment among many Japanese, as when everyone was sprayed with DDT, the reforms themselves were carefully and systematically planned with long-term vision.

Technological and Social Changes of Late Modernity

After the Occupation ended, the Japanese economy began to develop very rapidly. The gross national income doubled every five years

between 1955 and 1980. Refrigerators, washing machines, and television sets were present in almost all households by the 1960s. High school enrollment rose to 69.6 percent for women and 76.7 percent for men in 1965. University (including two year college) enrollment also increased: from 5 percent to 32.4 percent for women, and from 15 percent to 42.7 percent for men between 1955 and 1975.

The new Constitution proclaimed the fundamental rights of individuals, equality of men and women, and marriage based only on the voluntary consent of couples. In 1947, women's suffrage was approved. Also in that year the Civil Code was fundamentally reformed: equal rights for men and women were assumed for property, inheritance, custody of children, divorce, and other issues. The extended family system was legally abolished, and a democratic nuclear family was advocated. The average number of children per family decreased from four to two between 1950 and 1955. The Maternal and Child Health Act was promulgated in order to extend the scope of government involvement from simple child welfare to a systematic provision of a supportive environment for mothers and children. Between 1955 and 1965 the percentage of home deliveries dropped from 82.4 percent (93.4 percent in rural areas) to 16 percent (32.2 percent in rural areas). In 1950, 5.2 percent of all births took place under a doctor's supervision. By 1975, that number had risen to more than 90 percent, although changes in rural areas were slower. After the 1950s, medical technology, including antibiotics and the fetal monitor, became available. The neonatal intensive care unit was established in the late 1970s, followed by the perinatal intensive care unit in the mid 1980s.²⁰

Although these institutional and technological changes certainly contributed to a reduction in the IMR, their overall effect has probably been considerably less than the socioeconomic changes which took place over the same time span. Perhaps most important has been the persistence of a widely shared value system, the belief that an investment by both the government and families in infant and child health and well-being is of vital importance.

THE DARK SIDE OF IDEOLOGY

Historical analysis reveals that both the extended, patriarchal family system of the Meiji era and the militaristic nationalism of early

Showa affected the health of children in a positive way because high quality human capital was essential for the survival of these systems. Under the democratic mood of postwar Japan, the modern system of the welfare state was constructed; although modifications were made, prewar thinking continues to shape many current ideas and practices, including an ideology of motherhood, a social order in which the family system, and not the individual, is recognized as the basic unit, and a type of "nationalistic paternalism." The idea that a woman and her child are very close and cannot be conceptualized as independent individuals is commonly accepted. A mother bears the total responsibility for the well-being of her child. However, children are understood as belonging not only to their mothers but also to the family and society. In theory, mothers are protected and supported by society and by their own families as long as they stay in their maternal role. This situation probably contributed to the improvement of infant mortality rates in Japan, but not without some cost to certain individuals.

An Ideology of Motherhood

Perhaps the most striking feature of Japanese culture which affects attitudes towards MCH is the special emphasis given to "motherhood."²¹ With the recognition of the importance of the social position of women for health improvement, some researchers have questioned why infant mortality is low in Japan despite women's subordinate position in society. The percentage of women who are diet members (6.8 percent), executives (3.6 percent), tenured professors (6.2 percent), or public officials in management positions (0.7 percent) remains significantly lower than in other countries. However, while the social status of women may be low, the status of "mother" is very high. As long as a woman becomes a mother she is valued and respected in society.²² One survey showed that 65.2 percent of women and 69 percent of men think that "women should put their immediate families before themselves," and that women who put their own interests first are morally weak. Furthermore, MCH support has been organized on the assumption that mothers do not work outside the home.

In early modern Japan, women were expected to produce a successor for the family. Even today, childless women are regarded as less than full adults.²³ Adoption of unwanted children of relatives

was once common, but this is not usually an acceptable alternative today, nor is adoption of orphaned children; continuity of the bloodline is still valued and women feel anxious about raising children with whom they have no "natural" biological bond. Extensive use is therefore made of reproductive technologies such as in vitro fertilization.

The gender roles imposed by the state in the Meiji period intensified in postwar Japan when rapid economic development became the major goal of the nation. Japanese husbands have an international reputation for being "corporate soldiers," devoting all their time to work. Long hours and competition among colleagues, culminating at times in *karōshi* (death from overwork), make it almost impossible for husbands to assist a pregnant wife or to participate in child care. Family finances are managed by women in 70 percent of households, and women make most of the choices in connection with daily life. However, "important" decisions, such as the buying and selling of real estate, are still made mainly by the husband.

Another reason for the persistence of gender roles is discrimination against women in the workplace. On average, women earn only 61.5 percent of the wages of men, and with age the discrepancy becomes worse. Lower pay for women may be justified by the disruption of work due to childbirth and care, but a disparity exists even if an adjustment is made for educational level and type of work. An Equal Opportunity Law enacted in 1986 is rarely properly enforced.

If women continue to work after having children, they face many problems. It is very difficult to find reliable child-care facilities. Financial security during maternity leave and child-care leave, flexibility of working hours, understanding colleagues in the workplace, and paternal help are also recognized as being in short supply.

It is assumed by virtually everyone that husbands provide the financial support for their families. An allowance for dependents, including wives, is paid by most companies and many factories. This allowance will be made even if the wife is working, unless she earns an annual salary of one million yen or more. Women are thus indirectly encouraged to be "full time" or "professional" housewives (*sengyō shufu*), or else to be content with part-time and poorly paid work with no social security. These working women

who are the first to be fired when the economy goes into decline are an essential part of the Japanese labor force that remains hidden behind the dominant Japanese system of lifelong employment.²⁴

An Ideology of the Family

Japan's very low rate of out-of-wedlock births is easily explained.²⁵ When an unmarried woman becomes pregnant, the family pressures her to abort.²⁶ Alternatively, marriage is arranged before the birth, or the baby may be secretly adopted, institutionalized, or even registered as belonging to another family member. To most Japanese, raising a child without a father is an immoral act that disrupts the social order, brings shame to the family, and provides a bad environment for the child. There is also institutionalized discrimination against illegitimate children. Illegitimacy appears on official identity records (*koseki* and *jūminhyō*), which are used, for example, when seeking employment. Until recently, when it was recognized to be unconstitutional, the law which regulates inheritance deemed that an illegitimate child was only entitled to half what a legitimate child was entitled to.

The income of a family composed of a mother and her dependent children is about half the income of the average nuclear family. Although there is a specific law (*boshi oyobi kafu fukushi hō*) to support them, the majority of single mothers live in very poor financial situations, and, as we have seen, finding employment is not easy.

The idea of motherhood in Japan is socially constructed. Japan's population policy regulates the number of abortions performed and shapes values regulating which pregnancies should be terminated. A significant number of abortions are "socially high risk pregnancies," including out-of-wedlock and teenage pregnancies. Whereas the number of abortions among married women has declined, that among teenagers has increased quite significantly over the past few years. Abortion is described as a woman's "choice." Therefore the moral responsibility for terminating a pregnancy is placed entirely on her. Use of the Pill as a method of birth control remains illegal in Japan, partly due to the fact that the gynecological profession profits from doing abortions. It must be pointed out that many Japanese women profess a reluctance to take the Pill because they fear unwanted side effects. Legalization of the Pill in conjunction

with better sex education could reduce teenage pregnancies, but this problem is so stigmatized that it remains, as does AIDS, largely hidden and poorly managed.

Nationalism and the Low Birth Rate

From the Meiji period until World War II, children were thought of as the nation's treasure. However, in postwar Japan, the rhetoric about MCH has become democratic and humanistic. The well-being of mother and child is respected in its own right. When the total fertility rate dropped to 1.57 in 1989, shock and concern were expressed by politicians and leaders of business and industry. They were concerned about how the aging society would be supported by the decreasing number of young people. The Ministry of Health and Welfare started a "welcome baby campaign" highlighted by a song performed by musicians popular with younger women, and local governments planned similar events. A Child Care Leave Act was passed hastily in 1992, but the birth rate is still declining: 1.54 in 1990, and 1.45 in 1993. Currently the Ministry is planning a ten year project, known as the "Angel Plan," to establish a support system to improve child care.

The main cause of this decline in birth rate is postponement of marriage by women (the percentage of unmarried women in their late twenties increased from 20 percent to 40.2 percent between 1975 and 1990), although the average number of children produced by a married couple has not declined. Nevertheless, the somewhat emotional response to this "crisis" revealed the extent of nationalistic thinking about reproduction. Needless to say, the target for criticism was women. The leader of *Keidanren* (Federation of Economic Organizations), for example, has accused women of being selfish and simply enjoying their lives instead of having children. The mass media recently portrayed the low birth rate as "women's rebellion" or "*Shussan sutoraiki*" (childbirth strike). In addition, criticism was frequently expressed about women obtaining higher education.

Foreigners and the Ideology of Blood

Another dimension of nationalism is a rejection of "foreign blood" and an effort to maintain social stability by keeping a homogenous population. Children born in Japan do not receive Japanese nation-

ality if both parents are non-Japanese or if a Japanese father refuses to acknowledge paternity of a child born to a non-Japanese woman. The number of foreigners in Japan has increased significantly in the past five years due mainly to the shrinking local work force and to the demands of the entertainment industry. Health problems of foreigners have become a cause for concern, especially since the Ministry of Health and Welfare limited foreigners' rights under the Livelihood Protection Law in 1990. So far, the IMR of non-Japanese remains low (5.78 in 1991, 6.58 in 1992), but this figure includes only officially accepted foreigners. Considering the influence of both socioeconomic factors and lack of preventive care, a higher IMR can be expected among illegal immigrants. Concern about the low birth rate apparently does not encourage the idea of accepting more immigrants or protecting the children of immigrants, even when born in Japan. The welcome baby campaign is not for all babies.

Authority and Dependency

Major changes in the Japanese social system were brought about either from the top down or from the outside. Individual initiative has rarely been encouraged in Japan, and citizens are accustomed to, and even expect, protective management from above. Paternalistic policies have been justified through an emphasis on risk and early intervention; various health checks and a comprehensive screening program have been implemented without concern for cost-utility assessment.²⁷

Control of sexuality and reproduction by those in power is rarely examined critically. During the war, women were encouraged to produce as many children as possible, but shortly after the war women were urged to have only two or three children, and the number of abortions rose accordingly. Today, concerned about the population not replacing itself, the government is again actively encouraging women to have children. A recent official excuse for not putting the Pill on the market is that people will make less use of condoms, leading to an increase in AIDS.

Japan has often been described by social scientists as a culture which promotes dependence, particularly in situations where power is clearly unequally distributed. However, dependence does not imply, as North Americans might surmise, passivity. In Japan,

human relationships are usually understood as ongoing, reciprocal, mutually beneficial engagements. Even when power is unequally distributed, self-discipline and cooperation are expected on both sides. Care of mothers and children is particularly suited to this kind of societal arrangement. The government and families provide care for pregnant women provided that the women take care of their own health and that of their children in the prescribed manner.

CONCLUSION

Good health indicators, evidence of prosperity, a good social system, a sound economy, and a wise government have been a source of pride in Japan. The health-care system in Japan is, by all measures, effective, but if widely shared values, including a rejection of "foreign blood" and pressures on women for conformity with respect to motherhood, are integral to health-care policies, is this a good model for other countries? Should the IMR of a nation be reduced and kept low largely by means of policies which are nationalistic and which severely constrain the ways in which women can shape their own lives? If not, then it can no longer be assumed that women are "naturally" destined to devote their lives entirely to motherhood and family nurturance, nor should it be assumed that the family, extended or nuclear, will necessarily survive as a stable unit. The big questions now are how much longer women will allow their reproductive function to be monitored, how much longer they will provide the intense devotion and extensive time required to get their children through the Japanese school systems, and how much longer they will care for ailing elderly relatives. Many middle-aged and older Japanese women believe that younger generations have neither the stamina nor political commitment to bring about radical change. If they are correct, perhaps the status quo will be preserved a little longer.

ACKNOWLEDGMENTS

We appreciate the assistance of Drs. Tadamichi Meguro, Tadahiko Shimizu, Akira Shimouchi, Etsuji Okamoto, and Shizuko Sasaki with the collection of information.

The first author thanks Dr. Howard Hiatt for his course on the allocation of health resources offered at the Harvard School of Public Health in 1991.

ENDNOTES

- ¹John R. Evans, Karen Lashman, and Jeremy Warford, "Health Care in the Developing World: Problems of Scarcity and Choice," *New England Journal of Medicine* 305 (19) (1981): 11,717-727.
- ²J. C. Caldwell, "Routes to Low Mortality in Poor Countries," *Population Development Review* 12 (1986): 171-220. See also Peter Sandiford, Patricia Morales, Anna Gorter, Edward Coyle, and George Davey Smith, "Why Do Child Mortality Rates Fall: An Analysis of the Nicaraguan Experience," *American Journal of Public Health* 81 (1991): 30-37.
- ³Jann L. Murray and Merton Bernfield, "The Differential Effect of Prenatal Care on the Incidence of Low Birth Weight Among Blacks and Whites in a Prepaid Health Care Plan," *New England Journal of Medicine* 319 (21) (1988): 1385-391. William S. Nersesian, "Infant Mortality in Socially Vulnerable Populations," *Annual Review of Public Health* 9 (1988): 361-77. National Center for Health Statistics, *Proceedings of the International Collaborative Effort on Perinatal and Infant Mortality*, vol. 2 (Washington, D.C.: US Department of Health and Human Services, 1990).
- ⁴Health statistics used in this article come from the Ministry of Health and Welfare's population and vital statistics of 1992, unless otherwise mentioned.
- ⁵Ministry of Health and Welfare, *Boshi Eisei no Omonaru Tôkei* (Tokyo: Ministry of Health and Welfare, 1991).
- ⁶Ministry of Health and Welfare, *Shinshin shôgai kenkyû: Kôreika shakai o mukaeru ni atate no boshi hoken jigyo sakutei ni kansuru kenkyû hokoku* (Tokyo: Ministry of Health and Welfare, 1992), 11-31.
- ⁷Etsuko Matsuoka, "Media ni mirareru gendai nihon no ninshin, shussan" ("Pregnancy and Delivery in the Media of Modern Japan"), *Medical Humanity* 4 (2) (1989): 74-81 and Hiroyuki Ito, "Ninshin Shussan Jôhō no Hanran" ("A Flood of Information About Pregnancy and Child Birth"), *Boshi Hoken Jôhō* 20 (1989): 34-37.
- ⁸Emiko Ohnuki-Tierney, *Illness and Culture in Contemporary Japan* (Cambridge: Cambridge University Press, 1984), 181-88. See also Margaret Lock, *East Asian Medicine in Urban Japan: Varieties of Medical Experience* (Berkeley, Calif.: University of California Press, 1980), which discusses pluralism in the Japanese health-care system.
- ⁹The sources of polls and surveys related to women and work in this article are from the Ministry of Labor, *Chingin Kôzô Kihon Tôkei Chosa 1992 (Basic Statistic Survey on the Structure of Wages 1992)* (Tokyo: Ministry of Labor, 1992); Prime Minister's Office, *Danjo byôdô ni Kansuru Yoron Chôsa (Polls About Gender Equality)* (Tokyo: Prime Minister's Office, 1993); Ministry of Labor, *Hataraku josei no shûro to hoiku chôsa (Survey of Work and Child Care of Women)* (Tokyo: Ministry of Labor, 1993); and General Affairs Agency, *Hataraku Josei*

- no Jitsujō (Reality of Working Women)* (Tokyo: General Affairs Agency, 1993). Most of the results are also cited in the Prime Minister's Office, *Josei no Genjō to Seisaku (Current Condition of Women and Public Measures)* (Tokyo: Prime Minister's Office, 1993).
- ¹⁰Masako Horiguchi, "Yūshoku fujin no kakaeru mondaiten to sono taisaku" ("Problems and Measures in Connection with Working Women"), *Shussanki igaku* 24 (2) (1994): 244–48 and Junko Takashima, "Hataraku hahaoya e no kosodate shien no arikata" ("Ways to Support Child Care for the Working Mother"), *Shussanki igaku* 23 (6) (1993): 841–45.
- ¹¹Michael G. Marmot and George Davey-Smith, "Why are the Japanese Living Longer?," *British Medical Journal* 299 (1989): 1547–51; Economic Planning Agency, *Annual Report on National Accounts* (Tokyo: Economic Planning Agency, 1992).
- ¹²Shigeki Nishida, "Wagakuni kindai no shibō ritsu teika ni taisi iryōgijutsu ga hatashita yakuwari ni tsuite" ("The Role of Medical Technology in the Mortality Decline of Modern Japan"), *Journal of Public Health* 33 (1986): 529–33, 605–16.
- ¹³Sheldon Garon, "Rethinking Modernization in Japanese History," *The Journal of Asian Studies* 53 (1994): 346–66.
- ¹⁴For articles reviewing various aspects of women's lives, see: *Josei Shi Sōgō Kenkyū Kai Ed. Nihon Josei Seikatsu Shi (Japanese Women's Life Histories)*, vol. 4 Kindai and vol. 5 Gendai (Tokyo: Tokyo University Press, 1990). Articles by Kazuko Nagahara, Ryuichi Narita, and Mikako Sawayama in vol. 4 are especially relevant to women and health issues. See also, Chizuko Ueno, *Kindai Kazoku no Seiritsu to Shūen (Establishment and Closure of Modern Family)* (Tokyo: Iwanami, 1994). Official accounts on the history of public health in Japan are available in *Kōshū Eisei no Hatatsu* and *Zoku Kōshū Eisei no Hatatsu (Development of Public Health, vols. 1 and 2)* (Tokyo: Japan Public Health Association, 1967, 1983). Ministry of Health and Welfare, *Hokensho 30 nen shi (Thirty Years History of Public Health Centers)* (Tokyo: Nihon Kōshū Eisei Kyōkai, 1971).
- ¹⁵On *mabiki*, see Tokuji Chiba and Tadao Ōtsu, *Mabiki to Mizuko* (Tokyo: Nōsan Gyoson Bunka Kyōkai, 1983). Articles collected in *Iryō to Shakai, Gendai no Espwi* (Tokyo), 74, provide social and historical views on Japanese modern medicine.
- ¹⁶Carol Gluck, *Japan's Modern Myths: Ideology in the Late Meiji Period* (Princeton, N.J.: Princeton University Press, 1985).
- ¹⁷See the special issue of *Dædalus*, "Showa: The Japan of Hirohito," 119 (3) (Summer 1990), especially the articles by John W. Dower and Masakazu Yamazaki.
- ¹⁸Masanori Nakamura, *Keizai Hatten to Minshu Shūgi (Economic Development and Democracy)* (Tokyo: Iwanami, 1993), 118. This book also provides an excellent historical analysis on the development of Japan from the Meiji era until the present day.
- ¹⁹Ministry of Health and Welfare, *Yusei hogo tōkei shiryō* (Tokyo: Ministry of Health and Welfare, 1992).
- ²⁰Senya Toyama, Akiko Kitai, and Ryoji Takahara, "Kinnen no Nyūji Shibō no Henka ni tsuite," *Boshi Hoken Jōhō* 22 (1990): 5–16.

- ²¹On motherhood in Japan, see Bosei Kaidoku Kōza Eds, "*Bosei*" o *Kaidoku Suru* (Reading "Motherhood") (Tokyo: Yuhikaku, 1991).
- ²²Jungian psychologist Kawai Hayao claims that Japanese culture is based on *Bosei Genri* (The Motherhood Principle).
- ²³Atsumi Tsuge, "21 Seiki no Fufu to Kazoku" ("Married Couples and the Family in the 21st Century"), *Kôshûeisei* 56 (1992): 1–29.
- ²⁴Mary C. Brinton, *Women and the Economic Miracle: Gender and Work in Post-War Japan* (Berkeley, Calif.: University of California Press, 1993); Margaret Lock, *Encounters with Aging: Mythologies of Menopause in Japan and North America* (Berkeley, Calif.: University of California Press, 1993).
- ²⁵On the family in Japan, see the book series, *Henbô suru Kazoku* (*The Changing Family*) (Tokyo: Iwanami shoten, 1991).
- ²⁶An excellent analysis appears in Kyoko Yoshizumi, *Kongaishi no Shakaigaku* (*The Sociology of Children Born Out of Wedlock*) (Kyoto: Sekaishiso, 1993).
- ²⁷Naoki Ikegami, "Health Technology Development in Japan," *International Journal of Technology Assessment in Health Care* 4 (1989): 239–54.

Efficiency and Effectiveness in Health Care

JAPAN HAS BEEN ABLE TO ACHIEVE the seemingly impossible combination of cost containment, universal coverage, and no overt rationing of its health-care system. This has not been the result of pursuing effectiveness and efficiency, but of maintaining the status quo. The nationally uniform fee schedule has rewarded primary care services because it has been politically negotiated and because private practitioners continue to be the most powerful stakeholders. However, an increasing demand for long-term care, changing expectations of the consumer, and structural changes within the medical profession seem to require a different strategy.

HARNESSING HEALTH CARE'S UNIQUE CHARACTERISTICS

The Difficulty of Evaluating Its Effectiveness

The belief that health care is effective rests on the following assumptions: First, treatment is justified because it counteracts the pathological mechanism as it is understood by current biomedical research. Second, the improved condition of patients after treatment proves its effectiveness. And, third, this evidence has been validated by a clinical control study. Of the three assumptions, only the third provides conclusive evidence that the treatment is indeed better than a placebo, but very few procedures actually pass this test. Control studies take a great deal of time, and they can never keep up with the development of new technology and the physicians' eagerness to try new methods.¹ Even the few treatments which have

Naoki Ikegami is Professor of Health and Public Service Management at the Faculty of Policy Management and the School of Medicine, Keio University, Japan.

been rigorously evaluated are frequently used in situations where their effectiveness has not been clearly established, as witnessed by the expanded use of Cimetidine.² The claim has also been made that improvements in technology have made certain treatments effective despite the inconclusive results of the original study. Since clinicians witness an improvement in the patient's condition following treatment, they will disbelieve any conflicting studies and place their faith in technology. Thus, it is likely that health care will continue to be justified on the basis of the first and second assumptions, which do not provide any conclusive evidence.

It is useful to bear in mind that health care is administered in a closed, individual encounter between the physician and the patient. Since it is closed, the information that could be collected for the purposes of evaluation is limited and subject to bias: it can only be provided by either the physician or the patient. The accuracy of this information may vary. Because physicians make their livelihood by providing care, they cannot be relied on, for the most part, to document their services and their effectiveness accurately. And patients are seldom in a position to evaluate the utility of the process objectively. Moreover, since health care is individualized, even accurate information will, strictly speaking, be valid only for the unique conditions in which the encounter took place.³ Thus, it is hardly surprising that it is so difficult to evaluate the effectiveness, let alone the efficiency, of health care.

A Built-in Mechanism Which Resists Change

Health care is remarkably resistant to change. Since it is extremely difficult to monitor the private process, the solution has been to ensure the quality of the service by educating and training people so that they are capable of providing a standard level of care. Furthermore, the complexity of health care has made it necessary to subdivide the profession into specialties. These two factors have combined to make the hospital a "professional bureaucracy,"⁴ where conflict is avoided by a mutual agreement not to encroach on another's domain of expertise. The rigidity in the division of labor is not limited to hospitals and physicians but is pervasive throughout the health-care system. Any changes, therefore, will be uniformly resisted because they will lead to a disruption of the existing order and will pose a threat to the livelihood of health-care professionals.

A second reason for health care's resistance to change is the strongly held belief that the same quality of care should be provided to all, regardless of the ability to pay or geographical location. It would seem that if equality is to be valued above all, then it becomes next to impossible to introduce any notion of competition. Moreover, even if adverse selection can be eliminated, there would still be legitimate competition based on price and quality. And, should this fact become public knowledge, it would be extremely difficult to guarantee an equal provision of care. For example, if it became known that a certain provider gave the highest quality of care at a high price, that difference would have to be paid by either the patient or by the public. If the patient paid the difference, the public would say that the system was unfair because it provided the best care only to the wealthiest patients. However, even the latter alternative would not conform to egalitarian principles because the public would be paying extra for the patient who had the good fortune to receive care from a high price, high quality provider. There would still be a problem even if the provider had both high quality and low prices because of the resulting concentration of patients. This inequality, in turn, could only be resolved by differential payment, queuing, or a lottery system. Of the three, the lottery may be the most egalitarian but, at the same time, operationally the most impracticable. Thus, it is necessary to let the public assume that every facility and physician is providing a satisfactory level of care. This claim is often made by professional organizations to protect their members' interests, allowing the proponents of equality to find powerful allies among health-care providers. Both would maintain that selection of a facility should be based only on the patient's convenience and affinity, not on the quality of care. But since this does not lead to real competition, it will be very difficult to change provider behavior.

Third, consumer satisfaction is strongly related to consumer expectations, especially in health care where real benefits are so difficult to evaluate. Patients tend to prefer what has been provided to them in the past or what conforms to their notion of the current social norm. Furthermore, because the current pattern of services is regarded as an entitlement, any revisions will be fiercely resisted. Patients will willingly join forces with physicians so that the existing

pattern of resource allocation will be maintained even if the benefits are found to be, at best, marginal.

A Strategy for Cost Containment

Since the effectiveness of health care is established in a relatively small sector of health care, a global budget can be introduced without necessarily imperiling quality, at least in developed countries where poverty and illness do not affect the majority of citizens. However, the greatest problem plaguing health policymakers is the continuous introduction of budget straining new technology. In this case, resistance to change may be a positive attribute: once a global budget is introduced, it will be possible to rely on the medical profession's built-in propensity to maintain the status quo. It follows that policy should be directed at bolstering this tendency so that changes, including cost escalation, can be kept to a minimum. The cooperation of the most powerful stakeholders could be gained by pointing out that any revisions in the allocation pattern are likely to be to their disadvantage. This could be done without too much public protest because, as noted, patients' expectations are geared to what is currently provided.

Such a policy has, in fact, been pursued by Japan for the past ten years. This was not a conscious decision on the part of the government but was more the result of the country's unique sociohistorical characteristics. Naturally, the maintenance of the status quo came at a significant social cost. The condition of Japan's health-care system contrasts sharply with the advanced technology found in other industries.⁵ The fee-for-service independent practitioner is still the dominant figure. Most hospitals are small and physician owned with very basic amenities for patients. Formal quality assurance programs are largely nonexistent and providers are very reluctant to disclose any information. Despite this situation, the health indices of infant mortality and life expectancy at birth are excellent. Moreover, there is no overt rationing and the per capita number of CT scanners and other medical technology is the highest in the world.⁶

THE STATE OF THE JAPANESE SYSTEM

Historical Background

The history of health care in Japan is unique in two aspects. First, although in 1883 medical licenses were awarded only to those who

had studied Western medicine, provisions were made that allowed the preexisting practitioners of Chinese medicine (and even their sons) to continue to provide care and dispense medication. This meant that the country continued to have an extensive, well-established network of medical practitioners which was readily available to the general population.⁷ Private practitioners comprised the majority of health-care providers in the prewar Japan Medical Association (JMA), established in 1916. Second, there is only fragmented historical evidence of religious or secular orders which have provided institutionalized care.⁸ The hospitals which were established as part of the national policy of rapid westernization adopted in 1868 were the first to provide institutional care. Most developed from the clinics of physicians who had the business acumen to expand their practices. Those established in the public sector were designed for teaching, treating the military, and controlling the spread of communicable diseases.

When the government first legislated social insurance for manual workers in 1922, the two most powerful stakeholders within the medical profession were the private practitioners and a very small number of academic physicians. The hospital based physicians were subordinate to the university clinical departmental chairs who had in fact assigned them to their positions. Under these circumstances, it was not surprising that the initial fee schedule, the point fee system, was largely for the payment of ambulatory services provided by the practitioners in the clinics. In fact, its basic unit (the point) was the inclusive price of one standard drug per day. As the population covered by social insurance expanded, the original fee schedule was adopted by other plans. Private payment continued for the wealthy and was an important source of revenue for the prestigious physicians.

The JMA, representing the interests of the private practitioners, has played a dominant role in the fee schedule negotiations from the very beginning and continues to do so today. Although the number of hospital based physicians has increased dramatically, their position is still one of subordination since they remain under the patronage of the university professor. This hierarchical structure of the medical profession has hindered the development of collegial specialty associations despite the increase in specialized training. Furthermore, hospitals have been divided into a small and presti-

gious public sector and a large private sector, each viewing the other with antagonism.

How the Fee Schedule is Negotiated

Negotiations for revisions in the fee schedule take place in the Central Social Insurance Medical Care Council of the Ministry of Health and Welfare. The council is comprised of eight representatives from the providers (five physicians, two dentists, and one pharmacist), eight from the payers (four from the insurers, including two from the government, and two each from management and labor), and four who represent public interests (three economists and one lawyer). All five physicians are nominated by the JMA. Therefore, it is only through the JMA that physicians and hospitals can voice any of their demands.

The fee schedule negotiations are a two stage process. First, the estimated allowable increase in the total health-care expenditure is negotiated by the Ministry of Finance and the Ministry of Health and Welfare. The survey on the financial state of hospitals and clinics (the latter being equivalent to a study on physicians' incomes), which is made every second year, is presented to provide an objective basis for the change. Second, the JMA and the Ministry of Health and Welfare determine the allocation of the funds, individually readjusting the fees for each of the thousands of procedures and drugs. Since virtually all the fees are changed periodically, the effect of readjustment is never the same for any individual provider. It is difficult therefore to say definitively who has benefited most from the revision. The complexity of this process has worked towards containing the conflict among the providers.

However, the main reason why conflict has been kept to a minimum lies in the cardinal principle of maintaining the status quo: continuous efforts have been made to keep constant the relative share of the pie between inpatient and outpatient care, hospitals and clinics, and among the various specialties. Should any major imbalance or area of grievance arise, it would be redressed in the next round of fee negotiations. This principle has worked to the advantage of the private practitioners because, originally, health care consisted of mostly outpatient care provided from their clinics. Surgical fees were low to start with and have not, in general, increased as much as consultation fees.⁹ The JMA has demanded

and obtained various measures, such as additional consultation fees for internal medicine and pediatrics, so that the private practitioners' share could be maintained. At the same time, the JMA and the government have collaborated to contain hospital costs by decreasing hospitalization fees as the length of the stay increased and by lowering outpatient consultation fees. By such measures, the government has achieved the goal of cost containment while the JMA has been able to protect the interests of the private practitioners, its major constituency.

The power of the private practitioners is also apparent at the regional level. The local medical associations nominate members to the peer review boards of the clearinghouses for claims which are established in each prefecture. This board reviews the detailed item-by-item claims sent by every health-care facility at the beginning of each month and denies payment if they judge that the numbers of procedures or drugs have been excessive. The actual number of claims denied is less than 1 percent. Since these denials tend to be for the high cost claims, the review does have some discernible effect. Also, the denials tend to have a sentinel effect on the providers because if they gain a reputation for excessive services, their future claims would be subject to closer examination. These measures have restrained aggressive revenue maximization and market expansion by hospitals. The local medical associations also wield power through prefectural health planning committees (in most cases, the committee's chair is the chair of the prefectural medical association). The establishment of new hospitals or the expansion of bed capacity has to be approved by these committees.

The Pricing of New Technology

This description of the fee schedule seems incompatible with the fact that people in Japan receive the most CAT scan studies and hemodialysis for chronic renal failure per capita in the world. The answer lies in the general policy regarding new technology. The initial fee for the new technology is usually set at a level which is only slightly above that of the nearest existing technology. This does not normally cover the full costs, but many providers will try to attract patients and physicians by making the technology available. At the same time, manufacturers will compete to offer a less expensive model that has all the essential features of the more

expensive one. This will, in turn, trigger a downturn in prices (the CT scanner is a classic example). It should also be noted that diagnostic equipment is usually not labor intensive, and equipment or manpower that is of poor quality is not likely to have a direct adverse effect on the patient. Both would be more problematic for invasive technology, which might explain why the per capita number of surgeries in Japan is about a third of that of the United States.¹⁰ In any event, the low pricing of new technology has made it possible to maintain the appearance of professional freedom and avoid accusations of rationing while at the same time containing costs.

The case of pharmaceuticals illustrates the difficulty of using the fee schedule to control the use of technology. When a new drug is approved, its price is set at a maximum level of 10 percent higher than the nearest preexisting product. However, the manufacturer's return is much greater than this figure suggests. Market competition among the manufacturers regularly drives their prices down. Health-care providers then purchase the pharmaceuticals for the lowest possible price. The price paid for pharmaceuticals by social insurance is periodically reduced following the results of a national survey, conducted by the government every other year, of the price paid for pharmaceuticals by health-care providers. Through this process, the prices paid by insurers for some drugs have been reduced by over half in the past ten years. New drugs are therefore lucrative for manufacturers even if their prices are only 10 percent higher than the retail price of a preexisting drug. The purchase price of the preexisting drug would have already been reduced by the time the new drug is introduced.

If the marginal benefit of using the new drug is justified by the marginal increase in price, then the present system of allowing competitive market forces to drive prices down would be ideal. However, it is very difficult to evaluate the effectiveness of technology, including drugs. The present system gives incentives to prescribe more drugs and to shift to new drugs with higher prices. This may be one of the reasons why Japan's pharmaceutical expenditures are the highest in the world despite the fact that the individual prices for brand name drugs are lower in Japan than in the United States.¹¹

STRUCTURAL PROBLEMS FACING JAPAN

The best option for Japan may appear to be the continuation of the present policy. However, several structural changes have made maintaining the status quo increasingly difficult. For example, demographic and social transitions have resulted in an increased demand for long-term care. The number of the elderly aged sixty-five and older has increased from 5 percent in 1960 to 12 percent in 1990 and is estimated to reach 25 percent in 2020. With over half of all married women working outside the home, the traditional family support has become increasingly difficult to provide. This has led to a massive increase in the institutionalization of the elderly: from 1 percent in 1960 to 6 percent in 1990. Three-quarters are in hospitals because of the lack of alternative forms of care.¹² The share of long-term care in total health-care expenditures has been compressed because of the need to maintain a balance among the providers. This has led not only to a poor quality of care, but also to significant out-of-pocket costs for the patient (67,000 yen according to a survey made by Ryu Niki¹³). Although these extra charges are legally prohibited, they are unofficially condoned.

Another example is the changing expectations of the consumer, especially the younger urban population. The globalization of the economy has led to an enhanced awareness of the world outside of Japan, which is beginning to affect even the parochialism in health care. Concepts such as informed consent and patients' rights have become familiar to the general population through the influence of the mass media. At the industrial level, the process of harmonization in the pharmaceutical industry has highlighted the need to have more rigorous clinical trials and a more open distribution system for drugs. These comparisons are nearly always made with the United States. While the US health-care system leaves much to be desired, it does provide an alternative for the growing dissatisfaction with the system in Japan. One manifestation of the disenchantment with private practitioners is the growing concentration of patients in large prestigious hospitals. Since these hospitals receive subsidies they have been able to maintain standards which are closer to those found in Europe and North America.

A third example of structural change is occurring within the medical profession. When Takemi became chairman of the JMA in

1957, private practitioners were in the majority and most were in their forties. Now private practitioners constitute only a third of the total and their average age has risen to sixty-two. While the JMA continues to be a politically powerful force, it is hardly surprising that its leadership lacks dynamism. Not only are young physicians less prepared to go into open practice because of the specialized training they have received, but the increase in the initial capital investment that is required to set up the practice has become a major hurdle, especially in urban areas.

Both the government and the leaders of the JMA are, to a certain extent, aware of these monumental changes. But so far their hands have been tied by the general policy of maintaining the status quo. Any radical changes would, at least in the short term, raise costs and place the private practitioner at a disadvantage. The former would be precluded for fiscal reasons if it were financed by social insurance, or for egalitarian principles if it came from private payment. It is worthwhile to mention that unlike many West European countries, Japan does not have any officially recognized private payment system. Extra-payment is made only to elite physicians in the premier medical centers. It would be difficult to open this black market and expand it, especially since this would require the JMA to disregard the interests of the private practitioners.

STRATEGIES FOR CHANGE

It would be next to impossible to find the "complete cure" for Japan's health-care system. There is not going to be a "magic bullet" solution because of the strong resistance to change which is exemplified by the mutual accommodation between the government and the JMA. Changing the pattern of health-care delivery means changing the way health-care professionals earn their living. Since health care is regarded as an entitlement, and since health-care professionals have autonomy in their decisions, this would be difficult to realize. Furthermore, the value placed on egalitarian principles precludes consumer choice and tends to foster a rigid, monolithic provision of services. While it is true that consumers have become less satisfied with the health-care system in Japan, they would, at the same time, like to retain the easy access to outpatient care and the low costs which the current system provides.

Given these circumstances, change seems to be possible only on an incremental basis in areas where it is budget neutral and does not put private practitioners at a disadvantage. I have proposed reforms in two areas which meet these conditions.¹⁴ The first is aimed at the acute care provided by large public hospitals. The present fee-for-service system would be replaced by an inclusive per diem rate for inpatient care and inclusive per visit fees for outpatient care which would be negotiated at the prefectural level for each hospital that opts for this new payment mechanism. These rates would initially be set at each hospital's average earnings per patient. However, in the next year they would be tied to the hospital's performance based on such indicators as: productivity (bed occupancy rate and average length of stay); the provision of tertiary care (the number of complicated surgical operations and tertiary emergency care); the establishment of networks with the private practitioners (the ratio of referrals among new outpatients); and patient satisfaction. The number of indicators to be selected, their weights, and their targets would be negotiated at the local prefectural level by an organization having a membership composition similar to that of the Central Social Insurance Medical Care Council.

This new form of payment for acute care would be budget neutral and could lead to a decrease in total costs because public hospitals would have less need for subsidies. Efficiency savings could be realized because there would no longer be a need to dispense medication and order laboratory tests in order to generate revenue. The JMA has already proposed the implementation of this reform in the report from its Health Policy Committee. From their standpoint, it has the merit of freezing the share of the public hospitals and making those hospitals responsible for referring patients back to private practitioners.

The second reform is aimed at long-term care. The government has already established an inclusive per diem rate which geriatric hospitals may opt for instead of the regular fee-for-service payment if certain staffing requirements are met. The problem with this method is that although it solves the problem of over-medication and the ordering of too many laboratory tests, it gives the incentive for hospitals to admit only "light care" cases. To deal with this, I have been conducting validation studies on RUG-III (Resource Utilization Group, Version III), a case-mix grouping system for long-

term care developed in the United States. In order to take the first steps towards ensuring quality, validation is also concurrently being made on the MDS (Minimum Data Set) and RAPs (Resident Assessment Protocols). The MDS will provide comprehensive assessment while the RAPs will be useful in drawing care plans. Both are virtually unexplored areas in Japan. While these measures would not necessarily add more resources to long-term care, they would contribute to its more efficient use and, by highlighting the inadequacies in the present system, may lead to more radical reforms.

High-end acute care and long-term care are the two areas in which the greatest cost escalation is likely to occur if the present system continues. Coincidentally, they are also the areas which are marginal to the interests of the private practitioners, except as competitors for scarce resources. Therein lies the greatest hope for reform. Once these changes are effected, the stage will be set for the next round of reforms.

ENDNOTES

¹John E. Wennberg, "Outcomes Research, Cost Containment, and the Fear of Health Rationing," *New England Journal of Medicine* 322 (1990): 1202-204.

²Wayne J. Hall, "Use of Cimetidine in Hospital Patients," *Canadian Medical Journal* 124 (1981): 1579-584.

³Naoki Ikegami, "Best Medical Practice: The Case of Japan," *International Journal of Health Planning and Management* 4 (1989): 181-95.

⁴Henry Mintzberg, *Structuring in Fives: Designing Effective Organizations* (Englewood Cliffs, N.J.: Prentice Hall, 1983), 189-214.

⁵John K. Iglehart, "Health Policy Report: Japan's Medical Care System," *New England Journal of Medicine* 319 (1988): 807-12, 1166-172.

⁶Naoki Ikegami, "The Economics of Health Care in Japan," *Science* 258 (1992): 614-18.

⁷At the time of the first national census in 1871, the number who gave medical practice as their occupation amounted to 86.6 per hundred thousand. Shouichi Fuse, *History of Doctors: Its Japanese Characteristics (Ishi no rekishi: Sono Nihonteki Tokusei)* (Tokyo: Chuou Kouronsha, 1979).

⁸Shizu Sakai, *History of Health Care in Japan (Nihon no Iryoushi)* (Tokyo: Tokyo Shoseki, 1982).

⁹Yasuo Takagi, "The Effectiveness of the Japan Medical Association in Maintaining the Financial Position of Private Practitioners," in Naoki Ikegami and John C.

Campbell, eds., *Controlling Costs in Japanese Health Care* (Ann Arbor, Mich.: Michigan University Press, forthcoming).

¹⁰Toshihiko Hasegawa, "A Comparative Study on the Hospital Admission Rate Between the US and Japan," in *Ibid.*

¹¹Naoki Ikegami, "Comparison of Long Term Care for the Elderly Between Japan and the US," in Ikegami and Campbell, eds., *Controlling Costs in Japanese Health Care*.

¹²*Ibid.*

¹³Ryu Niki, *Health Care and the Fee Schedule of the 90's (90 Nendai no Iryou to Shinryou Houshu)* (Tokyo: Keisou Shobou, 1992), 198–230.

¹⁴Naoki Ikegami, "A New Fee Schedule to Realize Functional Differentiation—Reconstructing the Fee Schedule and Subsidies" ("Kinou Bunka notameno Shinshihari Houshiki—Shinryou Houshu to Hojyokin no Saikouchiku o Mezashite"), *Shakai Hoken Jyunpou* 1800 (1993): 6–9 and 1801 (1993): 10–13.

Doctors and other health workers are even less able to assure happiness than they are to assure health. If work is unsatisfying to many in modern society, psychopharmacology is a toxic and inappropriate remedy for correcting the resulting tension and alienation. If unemployed workers are depressed, mental-health counseling may be a temporary source of comfort, but the only genuine solution is full employment. If children fail to thrive, child-guidance workers may diminish their misery, but they cannot guarantee their flowering in the midst of social disaster. As physicians, our daily practice with human ailments makes us aware of the extent to which problems of ill health flow from failures in our political, economic, and social institutions. The redesign of these institutions is the central challenge for the coming century, and gives the greatest promise for improving public health.

Leon Eisenberg, M.D.

From "The Search for Care"
Dædalus 106 (1) (Winter 1977)

How the Developed Countries Became Rich

I

IN ADDRESSING A TOPIC THIS AMBITIOUS, I have no alternative but to confine myself to focusing upon essentials. In doing this I must sidestep some important aspects of the story. If I were proceeding at a more leisurely pace, I would want to assert that the proper response to the question of “how the developed countries became rich” is to say that they became rich in different ways. In particular, the countries that developed later did so in distinctly different ways from the countries that were the first to develop. The earlier developers were, in a very serious sense, exploring terrain that was distinctly *terra incognita*. Follower countries, by contrast, were largely moving along a previously traveled path.¹

When the United States began to industrialize early in the nineteenth century, it did not need to invent steam engines, railroads, power looms, or elaborate metallurgical processing technologies. These had already been developed in Western Europe. This is not to suggest that follower countries found development easy. In fact, follower countries encountered difficulties that the early developers did not. Today, the advanced industrial countries continue to generate innovations that may have very deleterious consequences to countries that are just aspiring to industrialize.

The twentieth century provides numerous examples of the substitution of new products for the old ones which some less developed countries have been heavily dependent on: synthetic fibers for cotton and wool; plastics for leather and for some nonferrous metals; fixing nitrogen from the atmosphere rather than relying on

Nathan Rosenberg is Fairleigh S. Dickinson, Jr. Professor of Public Policy in the Department of Economics at Stanford University.

deposits of guano; synthetic rubber for natural rubber; synthetic detergents for vegetable oils in the manufacture of soap; optical fibers for copper wires. An economy without some command of advanced technologies may be highly vulnerable to sudden changes in demand generated by the development of more sophisticated technologies abroad. If the domestic capability to adjust is limited, the economic consequences may well be devastating. Of course, it must be remembered that technical change and economic growth in the industrial world created the demand and profitable market opportunities for many resource-intensive products in the first place. For example, the extraordinarily high income levels of some Middle Eastern oil exporting countries would have been inconceivable in the absence of the intense demand for petroleum that was created by Western industrial technology.

II

What factors, then, accounted for the accelerated rate of economic growth that the presently-developed countries achieved at earlier stages in their history? If one puts aside what I would regard as critical facilitating institutional changes, it is possible to identify common strategic components in the experiences of each of the successful industrializers.²

A useful way to begin an examination of these components is to recall an illuminating exercise undertaken by Moses Abramovitz and other economists in the mid-1950s.³ Neoclassical economics had, for many years, suggested that the long-term rise in per capita incomes could be understood as a process of mobilizing more inputs into productive activity, primarily through a high rate of capital accumulation that raised the capital/labor ratio. But when Abramovitz compared the near quadrupling of US per capita incomes between 1869 and 1953 with the increase in conventionally measured inputs of capital and labor he found that the rise in such inputs was in fact quite small. The rise in inputs could account for only about 15 percent of the measured growth in output. Other economists, such as Robert Solow, arrived at similar results, even though they employed different methodologies and measurement techniques. It appeared that approximately 85 percent of the ob-

served growth in per capita incomes—"the residual"—could not be accounted for by the growth in conventionally measured inputs.

One might ask why the main neoclassical tradition remained immersed for so long in an apparent "dogmatic slumber," but there is, in fact, an excellent reason that has not yet been fully appreciated. The measurement of economic growth itself required a conceptual framework and a methodology for aggregating the myriad of activities that make up the daily life of a large economy. A way of measuring economic output and its changes over time was, in fact, a critically important intellectual "invention" of the interwar years. It was this prior conceptual thinking, and the painstaking empirical research to which it gave rise, that eventually made the discovery of "the residual" possible.

Once it became clear that only a small portion of the growth in measured output could be accounted for by the growth in measured inputs, numerous attempts were made to sort out the respective contributions of technological change and other factors to long-term productivity growth. The most systematic of these attempts were the growth accounting exercises of Edward Denison.⁴ There is now broad agreement that long-term economic growth is more a matter of using resources productively than merely employing more resources. Beyond that, however, there is still no close consensus on the quantitative importance of the separate variables.

III

The reason for this failure to achieve a greater consensus goes to the heart of our present concerns about the industrialization process. Successful industrialization, where it has occurred, has involved the creation of what might be called a growth *system*, a highly interactive set of forces and institutions that mutually influence one another. Within such a system, it is impossible to separate out the contribution of each variable and then to add up these separate contributions. Rather, within this system changes in any one variable need to be recognized as significantly influencing the "payoff" from other variables.

Thus, the interactions between technological change and capital formation have been central to this growth system. The large size of "the residual" pointed forcefully toward technological change—an

unmeasured input—as the critical variable. Yet, technological change, *by itself*, can probably exercise only a very limited impact upon productivity improvement. This is so because most technological changes exercise their impact only when they have been embodied in the form of new and more productive capital equipment. Thus, capital formation is typically the necessary vehicle for determining the speed with which, and the range over which, new technologies generate economic benefits. *Ceteris paribus*, societies that generate high rates of capital formation will derive the benefits of technological change more rapidly, and to a greater extent, than will societies with lower rates of capital formation.

But even this formulation is obviously too simple. The incentive to invest is not independent of the availability of new technologies that hold out the prospect of raising the rate of return on capital formation. Indeed, in a world of unchanging technology, the returns to capital formation might be expected to fall quickly to very low levels, as Malthus and Ricardo once feared. In such a world, even high levels of investment will eventually make relatively small contributions to economic growth. A construction industry that relied exclusively upon shovels would soon discover that when all workers were adequately equipped additional shovels contributed very little to further growth in output. On the other hand, in a regime of technological change, the availability of wheelbarrows would be expected to raise both labor productivity and the rate of return on capital.

Thus, the economic payoffs to technological change and capital formation have been closely intertwined, with the behavior of each simultaneously determining the value of the other. Similar observations hold, within an industrial system, between the commitment of resources to education and training of the labor force and the rate of technological change. An economy experiencing rapid technological change requires a more highly educated labor force, because such a labor force can more readily make the necessary kinds of adaptations and adjustments. Conversely, in a world of absolutely unchanging technology, the purely economic payoff to further education is likely to be low.

But here again, causality has to be seen as running in both directions. It is grossly erroneous to think of technology as some exogenous variable to which the labor force merely adjusts. On the

contrary, an educated labor force will have a greater capacity for generating technological change, not merely adjusting to it. One powerful insight that has emerged in the wake of the discovery of "the residual" is that economists have, in the past, defined capital formation far too narrowly. There can be no doubt that certain kinds of expenditures, not previously regarded as capital formation, have been crucial to the growth process in industrializing economies. If capital formation is conceived of in a generic sense to include any kind of present resource use that will demonstrably raise productivity in the future, it would incorporate expenditures on research and development (R&D) as well as some fraction of expenditures on education, nutrition, and health.

A rise in the rate of capital formation has historically been associated with the industrialization process, and the timing and the extent of the increase have shown considerable variation among countries. It is also true, however, that the national income accounts of successful economies have systematically understated the proportion of their resources that have been devoted to investment. A redefinition is necessary, one that expands beyond the narrow notion of capital as consisting exclusively of tangible reproducible goods. One estimate by Simon Kuznets attempted to provide some adjustment for the discrepancy between the generic and the conventional definitions of capital formation. Kuznets' estimate suggests that in one expanded definition the share of the gross national product (GNP) devoted to gross capital formation would be twice as large as in the conventional definition (see *Table 1*).

TABLE 1. Shares of GNP Components According to Various Assumptions

	Private and Public Consumption	Gross Capital Formation
National accounts definition	77 %	23 %
Omitting intermediate products	70	30
Allowing for investment in man but excluding income foregone from capital formation and GNP	58	42
Adding income foregone in formal education and on-the-job training	53	47

Source: Moses Abramovitz, "Manpower, Capital and Technology," in Ivar Berg, ed., *Human Resources and Economic Welfare, Essays in Honor of Eli Ginzberg* (New York: Columbia University Press, 1972), 68-69, n. 1.

This estimate indirectly raises another significant issue concerning the interaction of the variables. It is common to point to net capital formation when relating present resource use to future growth in productivity, since net capital formation indicates the share of an economy's output that is devoted to increasing the stock of capital. However, if technological change is proceeding at a rapid pace, a "replacement process" will occur; new capital will replace older capital. Because this process is not, strictly speaking, an increase in capital stock, it would not count towards net capital formation. Under these circumstances, gross capital formation may well be a more accurate measure than net capital formation of the extent to which resources are being applied to productivity increasing purposes. In industrial societies that already possess large stocks of capital, the proportion of GNP devoted to replacement is also large. Therefore, net capital formation figures will substantially understate the share of the economy's current output that is being devoted to uses that will raise future output. Kuznets has estimated that, for the period from 1919 to 1948, the United States devoted a modest 6.8 percent of its net national product (NNP) to net capital formation, but at the same time devoted a very substantial 18.2 percent of its GNP to gross capital formation.⁵

Some significant interactions within the growth system of industrialization involve the demographic variable. The growth in per capita output and the new uses to which industrial societies devote their resources—including improved nutrition and housing and better public sanitation facilities—have regularly brought with them reduced levels of mortality and greater life expectancy. The improvement in life expectancy alone has, in turn, had a significant impact upon the willingness to commit resources to education. Insofar as education is regarded as an investment that will yield a return in the form of a higher future income stream, an improvement in life expectancy raises the prospective return on such investments by extending the period of time over which these higher incomes will be experienced. Conversely, a country with a low life expectancy is also one in which, *ceteris paribus*, the returns on investment in education will be low. As Abramovitz has pointed out:

The high level of mortality rates a century and more ago meant not only that infant mortality was high, but also that the chance of survival through adolescence was small and life expectancy thereafter

much shorter than it is today. Crippling morbidity was also more common. All these conditions effectively reduced the prospective rewards to investment in man. To spend eight or ten years in school beyond the age when earnings might otherwise begin, to forego these earnings, and to bear the other expenses of schooling was obviously less attractive when the remaining span of working life was, say, twenty-five years, on the average than it is today when it is forty or forty-five years. In the same way, the lengthening span of working life must have made people more ready to accept the risks and costs of seeking their fortunes in distant places and in new occupations.⁶

Thus, industrialization and rapid economic growth were not simply consequences of a slow but sure accumulation of capital and labor augmented by exogenous technological change. Rather, they were the product of a "growth system" that was far more interactive than is commonly recognized. The loosening of constraints in one area was likely subsequently to allow the loosening of constraints in other areas. Improvements in one place triggered greater benefits elsewhere. The cumulative effect of this interactive process has been a truly astonishing rise in productive capacity.

Moreover, these improvements in productivity may even be applied to areas of human life and social organization where we normally think of such concepts as "productivity" or "efficiency" as foreign and inappropriate. Considered from a narrowly economic point of view, industrial societies operate at a much higher level of "demographic efficiency" than less developed countries (LDCs). Reducing the death rate among the young is a goal of overwhelming importance for its own sake, but it also has a positive effect on economic performance. Countries which have high infant mortality rates generally try to offset the high death rates with high birthrates, thereby "wasting" large amounts of resources on people who do not live long enough to enter the labor force.

The notion of improved "demographic efficiency" raises issues where an economist is well advised to tread lightly. But it needs at least to be observed that there is no clear historical correlation between the rate of population growth and the rate of growth of per capita income among the industrialized countries. What can be stated positively, however, is that over the past two hundred years all successfully industrializing countries experienced an acceleration not only in their rate of growth of per capita income, but in their

rate of population growth as well.⁷ The emerging growth system in industrial countries reconciled accelerated population growth with an acceleration in the growth of living standards. Put another way, it appears as if the same forces that were at work raising material productivity were also at work reducing premature mortality.

IV

Within the framework of interacting forces that have made up the industrial growth system, the application of scientific knowledge and scientific methodology to a progressively widening range of productivity activities was preeminent. Looking at the long term, Kuznets was surely right in asserting that "The epochal innovation that distinguishes the modern economic epoch is the extended application of science to problems of economic production."⁸

In precisely what way has science played this role? It is commonly argued or implied that the greatest economic benefits of science flow from research going on at the scientific frontier, and the usual formulations focus on causal forces that run exclusively in a "downstream" direction, from science to technology.

In fact, the economic benefits of frontier scientific research have always been highly uncertain. Some fundamental scientific breakthroughs have provided "only" intellectual or aesthetic satisfaction. Others have yielded benefits only many decades after their discovery. It is quite possible, although the proposition cannot be readily tested, that the lag between the growth of scientific knowledge at the frontier and useful applications of that knowledge has declined in recent decades, partly as a result of institutional innovations that facilitate those applications. The huge growth in the twentieth century of industrial research laboratories that monitor frontier research within the university community and elsewhere and that increasingly participate in frontier research themselves has certainly tended to reduce the length of the lags.⁹

The major functions of such industrial labs are to identify significant innovations wherever they may occur and to provide the scientific and technological expertise that make rapid imitation possible. In fact, from the point of view of deriving the economic benefits of science, the critical issues revolve around not creativity and originality in basic science but the effectiveness of the institu-

tional framework for the application of scientific knowledge, wherever and whenever that knowledge may have been generated. There has been in place for some time a communications network that is capable of transmitting the findings of scientific research all over the world. Developing successful commercial applications of this knowledge, however, is far from simple and straightforward.

It may be useful here to recall Alfred North Whitehead's observation that "The greatest invention of the nineteenth century was the invention of the method of invention."¹⁰ Although the sentence just quoted is well-known, his subsequent observation is not, but deserves to be: "It is a great mistake to think that the bare scientific idea is the required invention, so that it has only to be picked up and used. An intense period of imaginative design lies between. One element in the new method is just the discovery of how to set about bridging the gap between the scientific ideas, and the ultimate product. It is a process of disciplined attack upon one difficulty after another."¹¹ Whitehead understood that invention involved the linking of new scientific knowledge to the world of artifacts. But he also understood that this linkage was not easily achieved.

What appears to matter more than the quality of a country's basic science is the extent to which the activities of the scientific community are responsive to the needs of the larger society. Much of the rhetoric of academic science, with its overwhelming emphasis on the autonomy of science, seems to conspire against a proper understanding of this role. But the fact of the matter is that, in the course of the twentieth century, and with varying degrees of success, industrial societies have created increasingly dense networks of connection between the conduct of scientific research and the articulated needs of the larger social system. Indeed, the industrial research laboratory is essentially an institutional innovation in which the scientific research agenda is largely shaped by the needs of industrial technology. The role of industrial scientists is to improve the performance and reliability of technology. In doing this, they not only make use of the existing stock of scientific knowledge, but they frequently add to that stock, occasionally making additions that are judged to be of Nobel Prize quality. This should not be surprising. Industrial activity provides unique platforms from which to observe classes of physical phenomena that are unlikely to be observed elsewhere: premature corrosion of an underwater cable,

unidentifiable sources of interference in electromagnetic communications systems, or extreme heat generated on the surface of an aircraft as it attains supersonic speeds.¹²

The industrial research laboratory, in short, has rendered science more and more an endogenous activity increasingly shaped by economic forces. Science itself has become gradually incorporated, in the course of the twentieth century, into the growth system that has propelled industrial societies along the growth path that they have followed in the course of the last two centuries. That growth system, in which technological change has always played a central role, is now reinforced by a powerful scientific research capability which has strong public and private components: public subsidy of science, especially of a more basic nature; higher educational institutions that perform much of society's basic research in addition to their older, established training functions; and a research capability in private industry, where scientific resources are specifically directed toward strengthening and expanding a firm's technological capabilities.

V

Economists have long struggled to clarify the connections between the organization and conduct of economic activity and the level of material well-being that this activity ultimately delivers to the population. Central to this intellectual struggle, as we have seen, has been a distinction between the portion of an economy's output that caters to the immediate needs of a population (consumption) and the portion that is withheld from such use and devoted to increasing the future productive capacity of the economy (investment).

It is now also gradually coming to be understood that the "straight-forward" notion of consumption is built upon something less than *terra firma* as well. Recent research by economic, demographic, and biomedical historians has brought an increasing awareness of the ways in which rising standards of material well-being may have a positive effect on the future capacity of the human agent to produce useful goods and services. A major, though highly elusive, component of the growth system is the impact of improved nutrition on not just health and life expectancy but also on the central economic variable of labor productivity and therefore on economic growth

itself. In this respect, the time-honored distinction between consumption and investment is breaking down.

Although some of his conclusions remain controversial, the demographic research of Thomas McKeown, which culminated in his 1976 book, *The Modern Rise of Population*, generated great interest in the connections between rising per capita incomes and the decline in mortality.¹³ Perhaps his most important finding was a negative one: the dramatic decline in mortality from infectious disease in England and Wales between 1838 and 1970 could not be primarily attributed to public health measures or medical interventions. McKeown pointed out that the great reductions in mortality from infectious diseases, with the exception of smallpox, occurred long before the availability of effective medical interventions. By far, the largest single reduction in urban mortality in England and Wales before 1900 came from reduced mortality from tuberculosis. This decline was particularly relevant to McKeown's case: because tuberculosis is not a waterborne disease, its decline cannot be attributed to public health measures that made clean supplies of water more generally available. McKeown's general conclusion was that the main determinant of population growth was nutritional improvement associated with rising levels of per capita income.¹⁴

Of course, it had long been understood that there was a close connection between deficient nutritional inputs and life expectancy. Malthus himself had emphasized the contribution of periodic famines to high levels of premature deaths. Subsequent generations of economic and demographic historians have expanded on this theme without inquiring carefully into the extent of its importance, or the *longer-term* consequences of chronic malnutrition. Little attention was given to how poor nutrition in the early years of life contributed to stunted development, which in turn raises the likelihood of chronic disease later in life. Nor had much systematic observation been undertaken to connect dietary inputs with the capacity to sustain the energy required for productive activity. Recent findings on these questions appear to require a drastic revision of traditional views, or at least a drastic reassessment of the relative importance of certain variables. In fact, these findings may require considerable reformulation of the entire process by which the developed countries became rich.

As Robert Fogel has pointed out:

We now know... that famines accounted for less than 4 percent of the premature mortality of Malthus's age, and that the excess mortality of the ultra-poor (the bottom fifth of society) accounted for another sixth of premature mortality. About two-thirds of all premature mortality in Malthus's time came from the part of society that Malthus viewed as productive and healthy. Yet by current standards, even persons in the top half of the income distribution in Britain during the eighteenth century were stunted and wasted, suffered far more extensively from chronic diseases at young adult and middle ages than is true today, and died 30 years sooner than today.¹⁵

Dietary improvements have not only extended the span of human life, they have significantly improved labor force participation rates, especially from the lowest quintile of the population, and made possible the expenditure of much greater human effort on the part of the entire working population. Indeed, Fogel advances the remarkable claim that "...the average efficiency of the human engine in Britain increased by about 60 percent between 1790 and 1980. The combined effect of the increase in dietary energy available for work, and of the increased human efficiency in transforming dietary energy into work output, appears to account for about 50 percent of the British economic growth since 1790."¹⁶

CONCLUSION

Much has been learned about "how the developed countries became rich," and that learning process has identified significant causal relationships, the importance of which had been previously neglected. The growth system that has generated development has also improved the capabilities of the human agent, which have, in turn, vastly strengthened the forward thrust toward development. The subject of economic development and the formulation of policies that will accelerate it now clearly need to be approached with a far stronger focus upon the causes and consequences of human development.¹⁷

ENDNOTES

¹⁵See the illuminating examination of these issues in Moses Abramovitz, "Catching Up, Forging Ahead, and Falling Behind," *Journal of Economic History* (June 1986): 385-406.

²For a detailed historical treatment of the role of institutional change, see Nathan Rosenberg and L. E. Birdzell, Jr., *How the West Grew Rich* (New York: Basic Books, 1986).

³Moses Abramovitz, "Resource and Output Trends in the United States since 1870," *American Economic Review Papers and Proceedings* (May 1956) and Robert Solow, "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics* (August 1957).

⁴Edward Denison, *Accounting for United States Economic Growth, 1929–1969* (Washington, D.C.: Brookings, 1974); *Accounting for Slower Economic Growth: The United States in the 1970s* (Washington, D.C.: Brookings, 1979); and *Trends in American Economic Growth, 1929–1982* (Washington, D.C.: Brookings, 1985).

⁵Simon Kuznets, "Long-term Changes in National Income of U.S.A. since 1870," *Income and Wealth*, vol. II (Cambridge: Bowes and Bowes, 1952).

⁶Moses Abramovitz, "Manpower, Capital and Technology," in Ivar Berg, ed., *Human Resources and Economic Welfare, Essays in Honor of Eli Ginzberg* (New York: Columbia University Press, 1972), 54–55. Abramovitz's calculations draw upon Simon Kuznets, *Modern Economic Growth* (New Haven, Conn.: Yale University Press, 1966), table 5.2.

⁷Kuznets, *Modern Economic Growth*, chap. 2.

⁸*Ibid.*, 9.

⁹*The Directory of American Research and Technology*, 1992, listed 11,934 organizations, i.e., "all known non-government facilities currently active in any commercially-applicable basic and applied research, including development of products and processes." *Directory of American Research and Technology*, 1992 (New Providence, N.J.: R. R. Bowker, 1991), vii.

¹⁰Alfred North Whitehead, *Science and the Modern World* (New York: The Macmillan Company, 1925), 98.

¹¹*Ibid.*

¹²For further development of this point, see Nathan Rosenberg, "How Exogenous is Science?," in Nathan Rosenberg, *Inside the Black Box* (New York: Cambridge University Press, 1982), chap. 7.

¹³Thomas McKeown, *The Modern Rise of Population* (New York: Academic Press, 1976).

¹⁴The chain of influences that limited the rate of population growth "... was broken during the eighteenth and nineteenth centuries when advances in agriculture brought an increase in food supplies. The improvement in nutrition which followed led to the decline of infectious diseases and to a reduction of mortality and growth of population." *Ibid.*, 162–63. One of the most serious difficulties with McKeown's analysis is his failure to account for the much higher death rates from infectious diseases in urban areas than in rural areas, in spite of the fact that per capita incomes were substantially higher in urban areas. For a careful criticism of McKeown's argument, see S. Ryan Johansson, "Food for Thought: Rhetoric and Reality in Mortality History," unpublished manuscript, Stanford University, April 1992.

¹⁵Robert Fogel, "Economic Growth, Population Theory, and Physiology: The Bearing of Long-term Processes on the Making of Economic Policy," National Bureau of Economic Research, Working Paper No. 4638, 16–17.

¹⁶*Ibid.*, 24. According to Fogel, around 1790 the bottom 10 percent of the French labor force "...lacked the energy for regular work and the next 10 percent had enough energy for less than 3 hours of light work daily (0.52 hours of heavy work). Although the English situation was somewhat better, the bottom 3 percent of its labor force lacked the energy for any work, but the balance of the bottom 20 percent had enough energy for about 6 hours of light work (1.09 hours of heavy work) each day." *Ibid.*, 6–7.

¹⁷For some valuable perspectives on this subject that stress the centrality of early childhood development, see J. Fraser Mustard and John Frank, "The Determinants of Health," Canadian Institute for Advanced Research, Publication No. 5, August 1991.

Economic Growth and Investment in Children

IT WAS THE KIND OF SCENE YOU EXPECT to witness only in the movies, something that happens to other people but not to you. I was giving a talk to a group of economists in Finland. Someone entered the lecture hall from the back and interrupted. There had been an accident at home. My daughter was in the hospital. She had been crossing the street on a pedestrian crosswalk, holding hands with my wife, when a teenage driver swerved around the stopped cars and struck her.

The subject of my talk was economic growth. The central theme was that our actions determine how fast the economy grows. Conventional economic reasoning, with its emphasis on scarcity, diminishing returns, and the limits to growth, typically invokes technological change as a kind of *deus ex machina* that can reconcile these results with the historical record of increasing standards of living. This type of reasoning encourages the belief that in the long run economic outcomes are determined by forces that are beyond human control. Because each one of us, acting alone, cannot have much impact on the course of events, this notion has an appealing plausibility. But together our actions have a great deal of impact. Societies that support innovation and investment have shown dramatic economic improvement; societies that do not, have not.

I argued that the historical record should inspire optimism. There is reason to expect we can continue to make the kinds of scientific and engineering discoveries that drove economic growth in the past. But on the plane trip back from Finland I began to consider a possibility that has worried me ever since. In the pursuit of individ-

Paul M. Romer is Professor of Economics at the University of California at Berkeley.

ual self-interest we may be exploiting the available opportunities in a way that leaves future generations less well-off. While resource scarcity and environmental limitations are real concerns, I believe that economic growth, as conventionally measured, and improvements in the quality of life should both be possible. My fear is that our generation may make choices that undermine both, even if we are successful in developing new technologies. Future historians may conclude that we did not invest enough in our children.

THE POTENTIAL FOR GROWTH

Much of the recent work on economic growth addresses the question of whether it is possible to sustain a pattern of steadily rising standards of living.¹ Economics became known as the dismal science because economic theory seemed to suggest a negative answer to that question. When this issue was first considered, agriculture was the most important economic activity. Land and labor were perceived to be the fundamental economic inputs. The central concept in economic analysis was the law of diminishing returns. If the number of workers increased while the quantity of land remained constant, output per worker would fall. Many classical economists had a hand in discovering the principle of diminishing returns in agriculture, but Thomas Malthus drew the most attention to its implications.² Increases in population would inevitably lead to less food per person. Population increases would therefore drive wages and output per person down to the bare minimum necessary for human survival.

By the time economists had worked out all of the implications of diminishing returns in agriculture, the principle seemed obsolete. Agricultural output grew faster than the classical economists had predicted. In the public and professional imagination, the locus of production shifted from the farm to the factory. In their formal models of economic activity, economists shifted the key nonhuman input in production from land to capital. The more optimistic prospect that emerged from this analysis was that output per worker could grow beyond the subsistence level as the stock of capital per worker increased. In this view, it was capital that suffered from diminishing returns. As the rate of return and profits fell, capital

accumulation would eventually come to a halt. Growth in wages and output per worker would then also stop.

This prediction was far less pessimistic than the vision of chronic malnutrition and premature death held out by the Malthusian model. Nevertheless, it failed to explain the accelerating growth in per capita income that took place during the eighteenth and nineteenth centuries. It also failed to explain why the classical economists had been wrong about prospects in agriculture. How had it been possible for output per agricultural worker to increase by so much? Allowing for capital accumulation did not answer the question. Capital goods, such as steel plows, could raise the productivity of labor, but the quantity of land remained fixed. Classical economic theory predicted that as the quantities of both capital and labor increased relative to a quantity of land, diminishing returns should still set in.

Economists who worried less about how to model the world mathematically and more about capturing the essence of the growth process have long recognized that capital accumulation was only part of the story. Fundamental changes in the nature of work were also taking place, changes that went beyond using additional quantities of a fixed menu of existing capital goods. Farmers were not just using more steel plows. They were planting new varieties of seed, experimenting with different horticultural practices, and using newly discovered machines, such as reapers. Workers in the factory were not using more existing hand tools. Rather, they were employing new methods of production, like assembly from interchangeable parts, and using new kinds of machinery, like powered lathes and grinders. Although returns in any existing activity or investment opportunity tended to decline as people did more and more of the same old thing, some underlying process kept opening up new work methods and new capital goods to invest in. Pessimists thought that these discoveries were accidents that could not be relied upon for long, but the process of change has continued and shows no sign of slowing down.

The part of the economics profession that relied on formal mathematical models gave this process a label—technological change—but it was slow to do for it what had been done for the other processes driving growth: 1) identify a fundamental input; 2) explain technological change as being the result of accumulation of

this input; and 3) treat production of this input as an important economic activity. When pressed, economists who looked at the big picture in economic policy tended to identify scientific knowledge as the input driving technological change. They assumed that science was produced by the government and made freely available to the private sector. But the underlying model of factory-based production supplemented by freely available science did not clarify how it was that science “in the air” could influence economic activity.

Although the image of the economy as a large factory persists, economists are confronted once again with a world that is changing before their eyes. The share of employment in the manufacturing sector has been falling for several decades, echoing the large decline of employment in agriculture that took place over the last century.³ Despite its prominence in professional and popular discussions, manufacturing never dominated economic activity to the extent that agricultural production once had. Moreover, experience suggests that formal models of economic activity that are motivated by the image of the factory can be an impediment to careful discussions of the pervasive effects that intangible inputs have in all productive activities.

In popular discussions, a new, more suggestive image of economic activity based on information processing is beginning to displace the factory model. As those closest to the digital revolution have emphasized, there are three generic types of inputs in this activity: hardware, software, and wetware. Hardware takes the place of physical capital, including natural resources like land and raw materials. Software includes a wide range of inputs: literal computer code, designs for new machines, operating instructions for those machines, scientific principles, folk wisdom, the routines followed in a firm, the literal and figurative recipes we use, and even the language we speak. Software includes all knowledge that has been codified. Its defining attribute is that it can be copied, communicated, and used simultaneously by many people. Wetware, the things that are stored in the “wet” computer of the human brain, consists of both the human capital identified by economists and the tacit knowledge emphasized by philosophers and cognitive scientists.

The roles of hardware, software, and wetware can be discerned in all areas of production and in activities that are far removed from

digital computing. For example, in the construction of a new city near Shanghai, the government of Singapore says that its primary responsibility is to supply the software needed to make the city run (i.e., all the routines and operating procedures that have been developed in Singapore). The hardware is the physical infrastructure of the city. And the wetware will be supplied by Chinese officials, trained in Singapore, who will run the legal, administrative, and regulatory bureaucracies.

In the United States, much of the improvement in agricultural productivity can be traced to improvements in software (i.e., new instructions about how to run a farm) that were developed through agricultural research and disseminated by the agricultural extension service. Because of software's unique capacity for simultaneous, multiple use by many people, an innovation in software can have a far-reaching impact on all aspects of life. It is no accident that many of the important transitions in human history are marked either by a change in the software instructions we use to manipulate raw materials (the Neolithic Revolution, the Iron and Bronze Ages, the age of mass production) or by changes in the available methods for copying and communicating software (printing with moveable type, telecommunications, digital information processing).

THE POTENTIAL FOR DISCOVERING NEW SOFTWARE

A crucial fact about software is that a large quantity of it can be discovered and implemented on very simple hardware systems. The ability to keep discovering new and better software is at the heart of growth. The basic economic intuition about diminishing returns holds with full force if the available software is held constant. For example, if we worked with a fixed set of machine designs, the benefits from additional capital accumulation would soon disappear. There are only so many steel plows and grinding machines that we can profitably use.

The same inherent diminishing returns apply to wetware. Imagine if during the Middle Ages the feudal lords had insisted on raising the average level of education in the work force to six, twelve, or twenty years. Holding constant the software (the things that were known and taught in the monasteries) this increase in human capital would have had little impact on economic output. A

farmer who educated his children and workers at the leading educational institutions of the day would not have been able to increase his output per acre by very much. It is not the dissemination of old information but the steady arrival of new bits of software that makes education ever more valuable and provides new opportunities for growth.

But what should give us confidence that the pessimists are wrong when they assert that there is nothing left to discover? Imagine that before you are one hundred jars, each containing one element from the periodic table. A simple calculation shows that by sampling from the jars and mixing the elements together, you can make 10^{30} different mixtures. There are actually far more possibilities than this calculation suggests because one can combine the elements under various pressures and temperatures. It is a mathematical certainty that no one has ever tested the properties of the vast majority of the feasible mixtures. In fact, if each of the 5 billion or so people on earth had tried one new mixture every second since the Big Bang, we would only have tried about 5 percent of the possibilities.

Another example is a deck of cards. There are about 10^{68} different ways to order a deck of cards. Strange as it may seem, it is very easy for each of us to hold a deck of cards that are arranged in a way that has never existed before. The vast majority of these orderings will be of little intrinsic interest. A small number will produce striking results.

For reasons that seem to spring from some deep feature of human psychology, we find it difficult to believe that so many possibilities could emerge from familiar objects such as decks of cards or collections of jars. We therefore tend to agree with the pessimist who argues, for example, that almost all of the good mixtures from the elements of the periodic table have already been found. But if we have tried only a vanishingly small fraction of the mixtures, how could we possibly have found all the good ones? In physical systems, this would be possible only if our scientific knowledge told us where to look right from the start. Our scientific capabilities are not nearly this sophisticated.

There is a subspecialty in chemistry that tries new mixtures in roughly the way described above. Scientists working in this area are guided by hunches and general principles, but they never know exactly what they will find when they create a new mixture. Special-

ists call this field exploratory synthesis. Theory snobs call it “shake-and-bake” chemistry.

A few years ago, scientists pursuing this line of investigation produced a ceramic made by heating under pressure a mixture of four elements: copper, barium, oxygen, and yttrium. The ceramic turned out to be a superconductor at temperatures far higher than anyone had previously thought possible. Since then, many closely related mixtures of elements have been investigated. Some of them are more promising than the original mixture. At this point, it is difficult to tell what the full impact of these discoveries will be and although many technical difficulties remain to be solved, the possibilities are exciting to contemplate.

We have not yet reached the ten year anniversary of the discovery of high temperature superconductors, yet some firms are ready to apply this technology in the design of extremely sensitive detectors. We are, however, in the very early stages of development in this area. Whether or not the practical outcomes from this particular class of high temperature superconductors live up to the most optimistic predictions, the basic point remains. There are many possible ways to rearrange the material world, and some of these rearranged configurations may prove to be far more valuable than the individual elements alone.

Because we have only limited abilities of prediction, there is no substitute for continued exploration. In the story of the high temperature superconductors, the “shake-and-bake” chemists got the last laugh. The theorists failed to predict what would happen when this ceramic was formed, existing theory having suggested that this kind of mixture would not yield a superconductor. Several years after the properties of these ceramics were first demonstrated, theorists are still trying to come up with an explanation.

THE ENVIRONMENTAL EFFECTS OF ECONOMIC GROWTH

I have been a determined optimist, working in a profession that is based on institutionalized pessimism. Nevertheless, I am less optimistic about prospects for the future than I once was. My concern is not the one expressed by some environmentalists, especially the ones who do not want anything anywhere to change. If change is equated with a reduction in the quality of life, then there is neces-

sarily a conflict between quality of life and economic growth, which is driven by change. But if economic growth is interpreted as a process that makes available new opportunities and choices, all we have to do to improve the quality of life is make wise choices. Under the right conditions, economic growth often leads to increases in the quality of life. For example, over the last couple of centuries, improved nutrition has caused profound improvements in human health, one of the most important determinants of the quality of life.⁴ Similarly, increases in income in the wealthiest countries of the world have recently been accompanied by serious efforts to clean up the environment.

As the scope of the environmental calamities that took place behind the Iron Curtain becomes more evident, it is clearer than ever that widespread environmental degradation is a political problem, not an economic one. It is a sign of governmental failure, not an inevitable result of economic growth. If a government creates the right regulatory and legal framework for economic growth, the quality of life should improve. However, if the wrong framework is put in place, new technologies that harm the environment may be developed and existing technologies may be misused. Thus, the underlying problem lies in the institutional framework and the incentives it creates, not in technological progress itself.

Given the right incentives, the market economy could develop new methods of production that are more friendly to the environment than those currently in use. To appreciate how much scope there is for improving our current methods of production, think of the milk cow as a sophisticated, environmentally friendly chemical refinery. It takes no workers to operate, freeing people up to do other tasks. It maintains the internal conditions necessary to react to extreme fluctuations in external conditions. It fixes any mechanical damage that it sustains. It uses abundantly available biomass as its principle input and converts this into a safe and useful chemical.

There are many things that we can do with a computer that can store up to a gigabyte of software instructions, about what is stored in the DNA of a cow. We are, however, far from discovering what all of these things are. On a given computer we are constrained by the inherent physical limitations of the hardware. For example, we will never be able to write a program that will make a computer grow a new computer. Nature has an advantage because the soft-

ware coded in DNA includes the instructions for the assembly of the hardware on which it will run.

In the last century, there was at least one person, Charles Babbage, who had some idea about the power that would be unleashed by a computing machine that could follow software instructions. He tried to build a working machine but was stymied by the limitations of the hardware technology of the time. Even Babbage would have had difficulty imagining where work on computing machines would lead: spreadsheet programs, word processors, database programs, phone dialers, calendar programs, voice recognition, etc. It would have been impossible for him to appreciate how the software coded in DNA controls the more sophisticated hardware in a cow, directing microscopic assembly operations that make a new cow out of water and grass, guiding the chemical reactions needed to convert these same inputs into milk, and supervising the regulatory mechanisms that keep a cow up and running. With all we have learned in the last hundred years, it is still difficult to comprehend the sophistication with which these tasks are accomplished.

We are now in a position analogous to that of Babbage. While we can see the potential applications that will come from having software control hardware, we do not yet have the ability to make hardware that can take advantage of these possibilities. So far, our computers do little more than print and display information. They control the fuel flow in some cars, the flight controls in some airplanes, the switches in our telephone system, and the operations of some machine tools. But compared to what is physically possible, compared to what nature already does in the cow, our hardware for transforming physical objects is still quite primitive. In the meantime, however, we can try to improve the capabilities of our existing manufacturing and chemical processing technologies. We can also tinker at the edges with what nature does, trying to take advantage of its huge library of existing software for micromanufacturing, microchemical processing, and real-time process control.

In the lifetime of our children, substantial advances may be made in methods of production. Economic activity in the future may have fewer unwelcome side effects. New bio-remediation technologies may even be developed for cleaning up some of the environmental problems that previous generations caused. Even though the technological possibilities are quite promising, the possibility remains

that individual self-interest, operating in the wrong social and institutional environment, could lead to outcomes that are bad for everyone. Fortunately, current trends all point in a positive direction.

INVESTMENT IN CHILDREN

If there is evidence that much remains to be discovered, and if it is possible that we can harness these discoveries in ways that truly increase the quality of life, I must explain why I have become less optimistic about the prospects for future generations.

My daughter's accident forced me to confront the choices each of us face. At many junctures, guided by a complicated set of motivations, we choose between devoting additional attention to our children and pursuing other goals like monetary gain and professional success. In making these decisions we are guided by a complex set of motivations. We enjoy material goods and therefore value the wealth that makes it possible to acquire them. We care about the current and future happiness of our children. We are also influenced by the opinions that others hold about us. This last influence, the desire for status or high regard within some group of peers, typically gets short shrift in economic analysis.⁵ But it is a very powerful motivation.

In the weeks following my daughter's accident, I realized that, as an economist, I am a member of a strange virtual community, emblematic of many of the developments taking place in society as a whole. We communicate by phone, fax, email, and mail. A few of us meet regularly in offices, but mostly we see each other at meetings and conferences. We know a great deal about each other's writings and professional status, but we know very little about each other's personal and family lives. We are like members of a village, but all of the usual social and cultural context is stripped away.

At some level, this point should have been obvious all along. Still, I was surprised that such a dramatic event could take place in my life and that so few of my colleagues knew anything about it. This experience also brought back to the surface discomfort that had arisen in the past. At various times, each of us has to make sacrifices in our professional careers for the sake of our families. At an abstract level, my colleagues could understand the constraints

my family placed on me (and I placed on them), but, not knowing my family, it was sometimes hard for them to appreciate the choices I made.

There are many factors contributing to a general reduction in the sense of community that may once have encouraged parents to invest more heavily in their children. Urban settings offer a degree of anonymity that is not available in smaller towns. Many people who grew up in smaller towns have described to me how strongly the perceptions of others influenced their lives. Several people described how they were frequently admonished to consider "what the neighbors would think."

One theory of the demographic transition, the shift from high to low fertility rates, describes the decision to have fewer children as a response to increased employment opportunities for women.⁶ It is possible that these other factors may also have been operating. Women may have entered the paid labor force not just because the rewards from work have grown, but also because the rewards from child rearing have diminished. Whatever the source, the changes have been dramatic. In 1960 in the United States, about 18 percent of mothers with children under the age of six worked outside the home. By 1990, the fraction had increased to 59 percent.⁷

In discussions of child rearing, many people are uncomfortable with the cold, calculating language of trade-offs and incentives. We would like to believe that we do everything possible to help our children and that we do it out of selfless devotion. But the unavoidable fact is that every day each of us must choose between paying more attention to our children and pursuing alternative activities that can also give us satisfaction.

In a world of trade-offs, it makes a difference if our social environment awards less prestige and fewer psychic rewards to helping our children. The most important lesson in economics, far more important than the law of diminishing returns, is that people respond to incentives, even if they are not fully aware of doing so. If activities within the household offer fewer rewards than do activities outside the household, then people will change how and where they spend their time. For many people, the changes in behavior may be small, but taken as a whole the net effect can be great.

This does not mean that people today are more selfish than were people in the past, or that today's parents are more irresponsible. What has changed is the social environment that influences our incentives. These changes may be driven, at least in part, by the generally celebrated technological developments which have encouraged urbanization. And urbanization appears to be encouraging the development of separate adult communities of work, family communities based on residence, and disjointed communities in which our children play and learn. These developments may be gradually but steadily shifting the nature of competition for status and high regard: readily observable outcomes will matter more; things that require detailed personal knowledge will matter less.

The usual economic theory of the demographic transition suggests that a reduction in the quantity of children should be accompanied by an increase in the quality of care. And total investment per child should be higher. In fact, there are many indications that the opposite is occurring. Out-of-wedlock births, with the resulting reduction in parental investment by fathers, have increased substantially.⁸ In the last few decades, performance in school has been stagnant or has decreased.⁹ Achievement by grade-school children in the United States now lags significantly behind that in many other countries.¹⁰ One is struck by the paradox in the wealthiest nation in the world of worsening circumstances for large numbers of children and a high rate of growth of per capita income.

CONCLUSION

If I am right about the changes in our social environment and the effects that they will have on parental choice, the implications are grim. A growing body of evidence suggests that what happens during the first few years of a child's life has long lasting effects on many aspects of emotional, physical, and cognitive well-being.¹¹ Simultaneously, changes in the marketplace are placing ever greater value on the emotional and cognitive skills that are likely to suffer from reduced parental investment at early ages. These skills will also suffer from reduced attention to academic achievement and social and moral development during the school years.

There will always be some families that are lucky enough to have strong links to an extended family that pays attention to parental

investment in children, that can afford to send their children to suburban or private schools, that have not suffered from the misfortunes that can tear a family apart. Children from these families will do very well in the economy of the future. But there will always be other families that do not have these advantages. In a social environment that offers fewer rewards for parents who invest in their children—that places less social stigma on parents that fail to invest—children from less fortunate families may suffer large reductions in lifetime income, health, and general well-being. For society as a whole, this may lead to increasing income inequality, a lower average level of skill in the work force, higher crime rates, and a larger fraction of the population that is dependent on public support.

ACKNOWLEDGMENTS

I would like to thank Victor R. Fuchs for his reactions and encouragement during several discussions about the issues raised in this essay.

ENDNOTES

¹For an overview of work in this area, see the Symposium on Economic Growth in the *Journal of Economic Perspectives* 8 (Winter 1994).

²Robert Malthus, *An Essay on the Principle of Population*, 1798, ed. Philip Appleman (New York: Norton, 1976).

³In 1890, more than 40 percent of the labor force in the United States was employed in agriculture; in 1990 it was less than 3 percent. In 1970, about 26 percent of the labor force worked in manufacturing; by 1990 it had fallen to 18 percent. US Bureau of the Census, *Historical Statistics of the United States: Colonial Times to 1970*, bicentennial ed. (Washington, D.C.: US Bureau of the Census, 1975). Also, US Bureau of the Census, *Statistical Abstract of the United States*, 112th ed. (Washington, D.C.: US Bureau of the Census, 1992).

⁴See Robert W. Fogel, "Economic Growth, Population Theory, and Physiology: The Bearing of Long Term Processes on the Making of Economic Policy," *American Economic Review* 84 (June 1994): 369–95.

⁵See Robert Frank, *Choosing the Right Pond* (Oxford: Oxford University Press, 1985), for an example of economic analysis that takes status seriously.

⁶See Gary Becker, *A Treatise on the Family*, enl. ed. (Cambridge, Mass.: Harvard University Press, 1991).

⁷US Bureau of the Census, *Statistical Abstract of the United States*.

⁸For a discussion of trends in single parenthood and several other indicators of investment in children, see Victor R. Fuchs and Diane M. Reklis, "America's Children: Economic Perspectives and Policy Options," *Science* 255 (January 1992): 41–46.

⁹Ibid.

¹⁰See Harold W. Stevenson and James W. Stigler, *The Learning Gap: Why Our Schools are Failing and What We Can Learn From Japanese and Chinese Education* (New York: Summit Books, 1992).

¹¹Carnegie Task Force on Meeting the Needs of Young Children, *Starting Points: Meeting the Needs of our Youngest Children* (New York: Carnegie Corporation of New York, 1994). See also Daniel P. Keating and J. Fraser Mustard, "Social and Economic Factors in Human Development," in Council on Social Development, *Family Security in Insecure Times* (Ottawa: Council on Social Development, 1993).

Mechanisms of Brain Development and Their Role in Health and Well-Being

THERE IS ABUNDANT EVIDENCE that the brain influences many organ systems within the body, including the cardiovascular system¹ and the immune system.² There is also evidence that the brain's influence on these nonneural systems can help to promote well-being and long life. An individual's probability of success and, indeed, many of the abilities and attributes that an individual manifests result from an interaction between heredity and environment. Studies have shown that identical twins, who have the same genetic makeup, share behavioral attributes as diverse as appreciation for particular forms of music, intellectual abilities, and even criminal tendencies. In addition, there is accumulating evidence that individual experience literally shapes the developing brain and subsequent behavior.

Beginning life as a single cell, the human organism undergoes an extraordinarily complicated process of development in which cells divide repeatedly, migrate to their appropriate positions within the embryo, and differentiate. Parallel to the addition of neurons, neural pathways, and synapses (representing connections between cells) in the developing brain, there is an important process of *elimination*, or sculpting away. Cell death plays a crucial and necessary role in creating the nervous system.³

Neural development proceeds by *overproducing* neurons, connections, and pathways. Thereafter, development mirrors evolution and only the fittest of these neurons, synapses, and pathways are

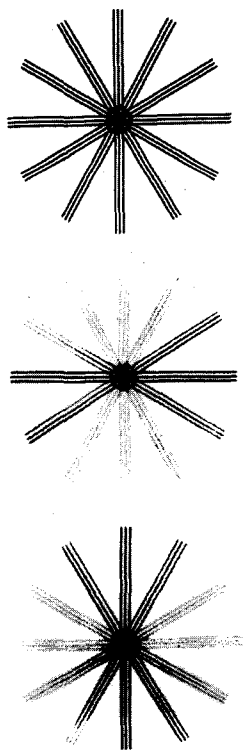
Max S. Cynader is Professor of Ophthalmology at The University of British Columbia, Vancouver, Canada, and a Fellow of The Canadian Institute for Advanced Research.

retained in the mature organism. Each neuron can be thought of as an individual, striving to grow, prosper, and accumulate as much synaptic territory and as many followers as possible. This type of biological aggression is valuable to neurons because it enables survival in spite of difficult developmental challenges. It seems intuitively reasonable that such a process should occur in a situation where many neurons will have to be eliminated.

SOME EXAMPLES FROM OUR SENSE OF SIGHT

Figure 1 shows a series of thin lines of different orientations. For the majority of viewers, these lines appear to be equally visible. Yet, for some individuals, the lines of a particular orientation are blurred and appear to be of lower contrast. The lower panels show how the lines in the top panel might appear to such individuals.

FIGURE 1.



Anomalous early visual experience results in permanent deficits in visual acuity for particular orientations.

This visual defect is frequently the result of uncorrected astigmatism, a clinical condition in which an individual's cornea is shaped like a lemon rather than like an orange.⁴ Because the corneal curvature is unequal, lines which are oriented in different directions cannot all be brought to focus at the same time. Hence, such individuals cannot see lines of a particular orientation with clarity. If this condition occurs in adults, an ophthalmologist or optometrist will prescribe a pair of glasses which will optically correct the corneal defect. However, if the astigmatism occurs in very young children and is not diagnosed and treated, the Darwinian competition that was described earlier takes place.

We know from studies of animal brains that we are born with a rudimentary set of cortical feature analyzers, each of which is involved with the analysis of lines of particular orientations. In the neonate, the initial set of rudimentary connections is set up independently of visual experience, and each set of orientation-specific analyzers comes to represent an equal area of the cortex. There are roughly equal amounts of cortical territory devoted to analysis of each of the different ranges of stimulus orientation. With early astigmatism, however, the amount of cortical territory devoted to the analysis of the orientation that was not seen well during early visual exposure shrinks substantially.

Exposure, or lack of exposure, to environmental stimuli can have a similar effect on the development of the cortex. For example, in one study two kittens were reared under environmental conditions which allowed a preponderance of either horizontal or vertical stimuli. The visual cortex of the kitten reared in an environment containing preponderantly *horizontal* stimuli responded to horizontal stimuli but not to vertical stimuli. The kitten reared in an environment with predominantly *vertical* stimuli responded to vertical stimuli but not to horizontal stimuli. Thus, in the cortex of the kitten reared in a predominantly horizontal environment, there were many more analyzers devoted to horizontal stimuli than to vertical stimuli. Just the opposite situation obtains in the kitten raised in the predominantly vertical environment. If one does not use a set of analyzers early in life, they will atrophy, and other competing analyzers will take over their portion of the cortex.

A condition such as early childhood astigmatism is frequently not recognized until a child goes to school and displays difficulty in

reading or in other visually demanding tasks. At this stage, the child may be prescribed glasses which provide a perfect compensation for the peripheral optical defect, but which may not be sufficient to cure the defect in the central nervous system. The use-dependent competitions depend on the exposure history of the organism, but only during a critical period early in postnatal development. During this critical period, the extent to which a particular region of the brain is used determines how it will function for the rest of the organism's life. If corrective measures, which counter the optical defect, are applied only after the critical period has passed, the distribution of feature analyzers will remain in a permanent imbalance.

Like the problem of astigmatism, amblyopia or "lazy eye"—when one eye, for some reason, does not provide a visual image which is as effective as that of the other eye, or if two eyes provide different and incongruent visual images at the same time—has a strong "critical period." If one eye fails to send the brain a useful message during early development, then its inputs are simply disconnected from the cortex. Extensive studies have delineated the timing of this critical period in both human and nonhuman species. The eventual visual acuity that is attained via a disused eye depends on the age of the organism and on the duration of the period of disuse.⁵ By disuse we do not necessarily mean blindness; reduced vision can result from cataract, or refractive errors such as myopia (nearsightedness) or hypermetropia (farsightedness).

What can be done about such disorders as early childhood astigmatism and amblyopia, which, when left untreated, cannot be corrected? First, we need better diagnostic approaches to identify these disorders early in life when therapy is still possible. In addition, neuroscientists worldwide are working to increase the plasticity and flexibility of the adult brain, to try to extend the critical period so that therapeutic approaches can be undertaken later in life. We have evidence that the critical period is not entirely fixed. It is, in fact, not really a period at all but rather a state of the brain. We know, for instance, that it is possible to extend the critical period in the visual cortex of animals well past its normal duration, simply by rearing animals in the dark. By *not using* the cortex throughout the naturally occurring critical period, it retains its capacity to be sculpted according to its eventual usage.⁶ This is not

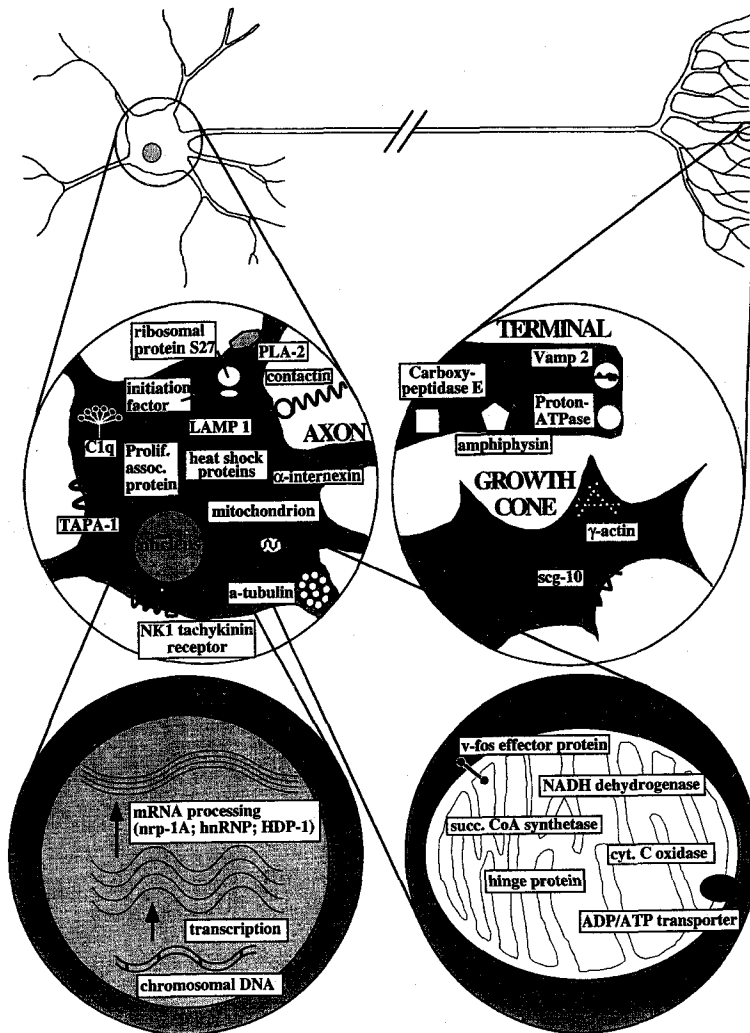
to imply that not using one's brain is advisable. Even though the visual cortex of an animal reared in the dark retains plasticity after the end of the naturally-occurring critical period, the functions of the cortex are compromised by this period of deprivation.⁷ Nonetheless, the finding that the critical period can be extended suggests that the critical period itself is not simply a biological clock that is ticking away as the animal ages. Rather, the use to which the brain has been put plays a role in determining its biochemical and physiological status, and hence its ability to learn and modify its circuitry according to usage.

Intensive efforts are now underway in many neuroscience laboratories to try to understand the essential nature of the brain state that allows for learning and modifiability. New molecular biological techniques have identified a series of genes that are turned on selectively in cortical neurons at the height of the critical period, and which are then turned off as the critical period wanes.⁸ *Figure 2* illustrates a prototypical neuron captured during its critical period. Each compartment of the neuron is shown in greater detail within one of the circles below the neuron.

At the middle left is a series of key molecules, including cell surface molecules such as contactin, TAPA-1, and the NK1 tachykinin receptors, which are overexpressed in cortical neurons during the critical period. In adult animals (beyond the critical period) these genes are shut off. The nucleus and mitochondrion are shown in the two circles at the bottom of *Figure 2*, illustrating a series of key genes that are overexpressed in these structures only during the critical period. The neuronal terminal (at the middle right) illustrates a series of genes and gene products that are turned on selectively within this compartment of the neuron during the critical period.

Neuroscientists are trying to recreate the critical period by using gene therapy to turn on certain genes. They are also trying to reinduce cortical plasticity by using therapies such as treatment with newly-discovered growth factors and other wonder molecules. In addition, behavioral strategies are being explored to try to extend the critical period into adulthood. One way to extend the critical period is not to use a particular piece of brain at all, which is not a very useful or practical approach. Another behavioral strategy is to *overutilize* the plasticity mechanisms during the normal critical

FIGURE 2.



Gene products that are over-expressed by neurons during the height of the critical period for activity-dependent brain plasticity.

Source: Shiv Prasad and Max Cynader, "Identification of cDNA Clones Expressed Selectively during the Critical Period for Visual Cortex Development by Subtractive Hybridization," *Brain Research* 639 (1994).

Reprinted with permission of Elsevier Science B.V.

period by subjecting a bit of cortex to continuous challenge. The theory is that by continuously exercising the plastic capabilities in a particular zone of cortex during the normally-occurring critical period, this plasticity might be prolonged. Just as exercising particular muscles strengthens them, so too exercising the plasticity mechanisms may strengthen the capability for plasticity.

THE GENERALITY OF CRITICAL PERIOD EFFECTS

Although the examples presented thus far were selected from the sense of vision, there is no doubt that use-dependent selection of particular pathways and neural circuits occurs in other sensory pathways as well, including hearing and the senses of smell and touch.⁹ There is also strong evidence for critical periods in the development of higher cognitive functions such as language processing. For instance, the guttural sound "ch" is an important component of several languages, including Japanese, Spanish, and German, but it does not play an important role in English. During the "babbling phase" of language development in the first few years of life, all children make and use this guttural sound, but children who grow up learning only English lose the ability to make this guttural sound.

There is abundant evidence that the best time to learn new languages is relatively early in life. Individuals who are exposed to a foreign language after they are eight or ten years old will eventually learn the foreign language, but they may never acquire the facility with this language that occurs naturally in young children. In fact, the very brain structures that are used to learn a second language are different from those involved in learning a mother tongue. An intriguing recent case report¹⁰ showed that a bilingual patient who suffered a specific stroke-like lesion developed severe impairments in mother tongue production, with substantially better performance in her rarely-spoken second language. This is consistent with the notion that learning languages is still possible in older individuals, but that the neural mechanisms and even the brain structures that are involved are quite different.

Other abilities also show a clear critical or sensitive period during development. They include complex motor skills such as riding a bike, skiing, or swimming. In addition, abstract intellectual abili-

ties, such as musical and mathematical capabilities, are strongly age-related. Emotional attachments too are tightly regulated during development. The quality of the developing organism's early social and maternal environment plays an important role in determining the individual's social responsivity throughout life.¹¹ A now famous series of experiments showed that depriving young monkeys of normal maternal support and stimulation during the first few weeks of postnatal life was sufficient to produce profound long-term social damage to those monkeys. Monkeys reared without their mothers were unable to have normal relationships, either social or sexual. And female monkeys reared without maternal support failed to show normal maternal behavior.¹²

Importantly, it appears as though these critical periods do not occur in all parts of the developing cortex at the same time. There are suggestions that different *layers* of the cortex may undergo critical periods at different times. Some cortical fields show evidence of retaining plasticity even up to adulthood. Some of the biochemical features that we have observed in the visual cortex at one particular age are in fact present in other cortical structures at other ages. Emerging brain imaging technologies such as functional magnetic resonance imaging (MRI), which enables visualization of the activity of particular pieces of brain during specific tasks, may enable physicians to direct particular training and exercise to those parts of the brain that are most ready to accept it at any given age.

STRESS AND CRITICAL PERIODS

In the presence of short-term stressors, the body mounts an organized defense, secreting a series of releasing factors and hormones which increase the heart rate, alter the processing of glucose, and dilate the pupils. While there is no question that the short-term response of the body to stress is beneficial and necessary, there is growing evidence that long-term stress and the body's response to it are counterproductive and deleterious.

Chronic stress can cause depression of immune function and other body systems that are controlled by the brain, such as the cardiovascular system. Furthermore, chronic stress has a deleterious effect on brain function. Repeated and prolonged exposure to stress hormones causes neurons to die.¹³ Interestingly, emerging evidence

suggests that those neurons that are *most ready to learn*, those that are within their critical periods, are the *most vulnerable* to the cell death and degeneration associated with chronic stress. This is consistent with other findings¹⁴ which show that social stressors can activate hormones that adversely affect brain functions, including learning and memory, and that these effects may be permanent. In addition, the stresses to which we are exposed early in life, during a critical period, may modify our ability to *modulate and control* responses to stressors later on in life.

SOCIETAL IMPLICATIONS

It is clear that the period of early postnatal development is a crucial one for the development of the neocortex, and hence for the sensory, cognitive, and social skills that will be necessary to guide the organism through life. This paper has concentrated on the deleterious consequences of an inappropriate environment, but there is a more positive side to the knowledge that we are gaining about brain development. If we can identify, assess, and evaluate early deficits within the sensory, cognitive, and emotional environment of the young child, then we can take steps to prevent damage.

The old nostrum, "An ounce of prevention is worth a pound of cure," applies in brain development as well as in other areas. We need to recognize the first few years of postnatal life as a key time for child development. We need to educate parents to recognize the importance of providing adequate social and cognitive stimulation for their children, and we need to educate our officials to support increased involvement of parents with their young children during this window of development. In contrast with many other industrialized nations, both the United States and Canada fail to give parents sufficient time to spend with their newborn children and are sadly deficient in ensuring postnatal health care for mothers and infants. We need to redress these deficiencies, and to emphasize education and support programs for families with infants and toddlers. The costs to society for neglecting the development of our young children are likely to be extreme.

ENDNOTES

- ¹Franco R. Calaresu and Douglas J. McKittrick, "Interaction of Putative Transmitters in Central Nervous Pathways Involved in the Control of Heart Rate and Arterial Pressure," *Cardioscience* 2 (1991): 147-54.
- ²Robert M. Sapolsky, *Stress, the Aging Brain, and the Mechanisms of Neuron Death* (Cambridge, Mass: MIT Press, 1992).
- ³More than half of the cells in the retina die before a child is one year of age and over one-third of the neurons in the cerebral cortex are eliminated in the first three years of postnatal life. Early in postnatal life there is a substantial pathway connecting the visual and auditory cortex, which largely disappears later in postnatal life. See Douglas L. Meinecke and Pasko Rakic, "Expression of GABA and GABA_A Receptors by Neurons of the Subplate Zone in Developing Primate Occipital Cortex: Evidence for Transient Local Circuits," *Journal of Comparative Neurology* 317 (1992): 91-101.
- ⁴Donald E. Mitchell, Ralph D. Freeman, Michel Millodot, and Gunella Haegerstrom, "Meridional Amblyopia: Evidence for Modification of the Human Visual System by Early Visual Experience," *Vision Research* 13 (1973): 535-58.
- ⁵Donald E. Mitchell and Brian Timney, "Postnatal Development of Function in the Mammalian Visual System," in I. Darian-Smith, ed., *Handbook of Physiology, Section I: The Nervous System*, vol. 3, pt. 1 (Bethesda, Md.: American Physiological Society, 1984), 507-55.
- ⁶Max Cynader and Donald E. Mitchell, "Prolonged Sensitivity to Monocular Deprivation in Dark-Reared Cats," *Journal of Neurophysiology* 43 (1980): 1026 - 40.
- ⁷Max Cynader, Nancy Berman, and Alan Hein, "Recovery of Function in Cat Visual Cortex Following Prolonged Deprivation," *Experimental Brain Research* 25 (1976): 139-56.
- ⁸Shiv Prasad and Max Cynader, "Identification of cDNA Clones Expressed Selectively During the Critical Period for Visual Cortex Development by Subtractive Hybridization," *Brain Research* 639 (1994): 73- 84.
- ⁹Peter C. Brunjes, "Unilateral Naris Closure and Olfactory System Development," *Brain Research Review* 19 (1994): 146 - 60.
- ¹⁰Salvatore Aglioti and Franco Fabbro, "Paradoxical Selective Recovery in a Bilingual Aphasic Following Subcortical Lesion," *Neuroreport* 4 (1993): 1359-362.
- ¹¹Stephen J. Suomi, "Social and Biological Pathways that Contribute to Variations in Health Status: Evidence from Primate Studies," in Canadian Institute for Advanced Research, *Prosperity, Health and Well-Being: Proceedings of the 11th Honda Foundation Discoveries Symposium* (Toronto: The Canadian Institute for Advanced Research, October 1993); R. Loeber, "Development and Risk Factors of Juvenile Antisocial Behavior and Delinquency," *Clinical Psychology Review* 10 (1990): 1- 41; K. MacDonald, "Early Experience, Relative Plasticity and Social Development," *Developmental Review* 5 (1985): 99-121.
- ¹²Harry F. Harlow, *Learning to Love* (San Francisco, Calif.: Albion Press, 1971).

¹³Subjecting animals to chronic stress (i.e., forcing them to endure inescapable shocks) results in cell death in various parts of the brain. See Sapolsky, *Stress, the Aging Brain, and the Mechanisms of Neuron Death*.

¹⁴See Bruce S. McEwen, *Hormones and Brain Development*, address to the American Health Foundation, Washington, D.C., April 1992.

In the course of this essay, we have noted several heavy burdens of distress and disability that have largely been neglected by science until the last quarter-century. One phase of the life cycle that is still being neglected in the field of mental health is the onset of adolescence, a time of lifelong importance in the formation of crucial behavior patterns involving habits that influence future health, such as cigarette smoking, alcohol and other drug use, automobile driving, nutrition and exercise, and human relationships. These extensive changes in behavior are associated with drastic physiological and biochemical changes in the body. Messages from the brain influence the pituitary gland to release hormones which in turn control the output of sex hormones. These sex hormones have far-reaching biochemical influences throughout the organism, including effects on cells and circuits in the brain. Although it is plausible to suppose that there are specific links between certain hormonal changes and concomitant changes in fundamental components of sexual, aggressive, and emotional behavior, virtually no well substantiated information exists to confirm this supposition, despite its clinical significance.

Philip Berger, M.D., Beatrix Hamburg, M.D.,
and David Hamburg, M.D.

From "Mental Health: Progress and Problems"
Dædalus 106 (1) (Winter 1977)

The Lifelong Impact of Childhood Experiences: A Population Health Perspective

IT IS GENERALLY BELIEVED THAT EARLY childhood experiences have a powerful effect on one's life. Many late twentieth-century institutions and orthodoxies about childhood development have been based on this belief. We try to protect the physical, social, and ideological environment of children (television violence notwithstanding), and we continue to invest heavily in education for the young. Yet, we have no working consensus as to how, and to what extent, the experiences of childhood influence later life.

A working consensus once existed. The movements to eliminate child labor and bring about universal, publicly funded education were sufficiently articulated by their advocates, understood by the population, and accepted by those with antagonistic material interests to overcome the prevailing norms of the time. The achievements of that era have endured and reflect themselves in beliefs we hold today. For instance, we still believe that having a decent education and growing up in a stimulating and safe environment improves a child's life chances.

The threats facing children today are no less significant than those of the past, but they fail to arouse the same indignation as, for example, the death and deprivation of so many children during the period of child labor. Across North America approximately 50 percent of single parent families live in poverty,¹ more than twice as many as in Western Europe. The number of children so affected is likely to increase with the rising number of families headed by a single parent and with economic stagnation. Among two parent/

Clyde Hertzman is Associate Professor, Department of Health Care and Epidemiology, The University of British Columbia and a Fellow of The Canadian Institute for Advanced Research, Canada.

two wage earner families, still the most common type in North America,² child-care arrangements are in a state of chaos.³ Children are often cared for by three to five different people each week.

No one suggests that poverty is preferable to affluence in raising children or that exhaustion makes for good parenting. Yet our agreement on these points has not translated into a working consensus. Many of the debates that currently dominate our attention lead away from the areas of social consensus and obscure the principal issues. Particularly unhelpful are those debates which focus on the legitimacy of specific family arrangements (i.e., the paternalistic two-parent family, gay parents, the role of divorce) because they detract our attention from significant new insights about the lifelong impact of childhood experiences, which could become the basis of a new consensus for improving children's life chances, regardless of their domestic arrangements.

WHAT MAKES SOME POPULATIONS HEALTHIER THAN OTHERS?

Even in a world without child labor and sweat shops the quality of the early childhood environment has a more profound and lasting effect on health, well-being, and competence than we have been prepared to recognize. This is the conclusion of recent investigations that have asked the question: what makes some populations healthier than others? Among the member countries of the Organization of Economic Cooperation and Development (OECD) the answer is simple: health status increases with increasing socioeconomic status.⁴ The pattern is not, typically, a simple difference between the healthy rich and the unhealthy poor. Rather, the health status of a particular class within the population seems to be better than the classes below it and worse than the classes above it. In other words, middle-class people may live longer and healthier lives than the poor, as we would expect, but they also live shorter, less healthy lives than the rich. This observation challenges our belief that the only relevant class factor is a level of poverty which would produce shortages of materials such as food, clothing, and shelter, which are necessary for the maintenance of good health. But there is more at stake.

The pattern described above is remarkably consistent across countries regardless of whether the classes are defined by levels of

income, education, or occupation; so much so that it has been canonized as “the socioeconomic gradient in health.”⁵ The pattern has remained essentially unchanged since the beginning of the twentieth century, despite the fact that the principal causes of death have changed completely. From the perspective of modern biomedicine and public health this is more than just counterintuitive; it challenges the role and function of specific disease labels. Ever since cigarette smoking was identified as a cause of lung cancer, we have become accustomed to fighting diseases one by one, identifying “risk factors” for particular diseases, and trying to eliminate them. However, if the pattern does not depend upon a specific set of diseases, then it is doubtful that “risk factors” such as smoking, drinking, and not wearing seat belts can account for the incidence of disease, even if these behaviors are more common among the poor and undereducated.

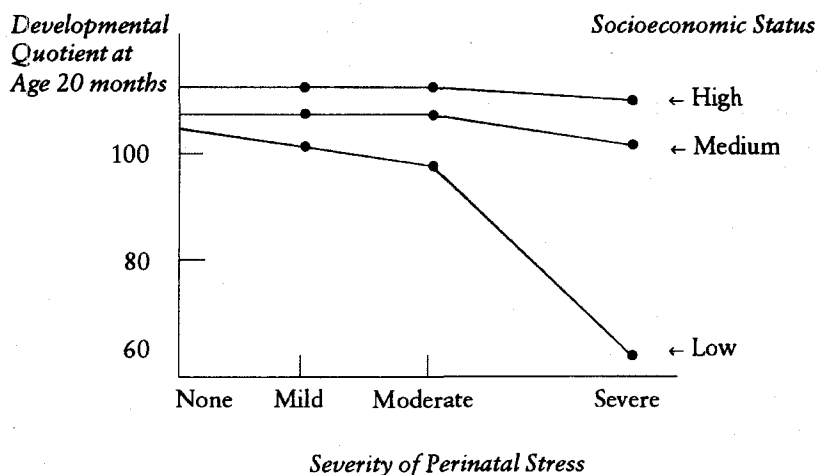
If negative health behaviors and limited access to health care are not the answer, then what is? The implication is that living conditions directly affect health. One’s place in the social hierarchy and the experiences which follow from it appear to be powerful determinants of the length and healthfulness of life. After all, the central nervous system, which is the center of human consciousness, “talks to” the immune, hormone, and clotting systems, which can, in turn, increase or decrease levels of resistance to disease and change the long-term function of vital organs of the body.⁶ But how does human experience initiate and sustain the processes which lead to differing levels of healthfulness?

Relevant research and insights which address this question are found in a bewildering variety of disciplines, levels of population aggregation (from the cell to the nation), and species of life. The following examples highlight different aspects of its complexity.

Example 1. High class status is a powerful buffer against both internal and environmental threats to successful human development. For example, it is well known that childbirth is a stressful process for both mother and baby and that long-term damage can occur as a result of birth traumas. Most newborns experience mild or moderate levels of stress, but if the stresses are severe, there are likely to be long-term effects. It turns out that lower-class children are much more vulnerable to these stresses than are upper-class children. *Figure 1*, from the Kauai birth cohort study,⁷ shows that

severe pregnancy, labor, and delivery complications did not lead to impairment of the physical and psychological development of children from upper-class families in the way that they did for lower-class children. This example suggests that high class status not only protects healthy children from future risks and vulnerabilities, but can actually help reverse the long-term consequences of birth stresses.

FIGURE 1. Interaction of Perinatal Stress and Socioeconomic Status in the Developmental Quotient

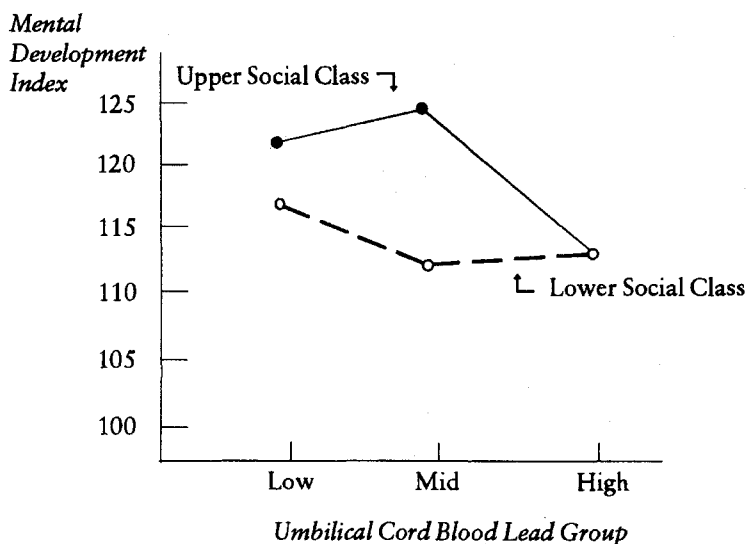


Source: Emmy E. Werner, "Children of the Garden Island," *Scientific American* (April 1989).

Figure 2 shows a remarkable parallel with exposure to lead in the environment. Lead exposure is a well-known cause of learning difficulties and nonadaptive behavior patterns in school-aged children.⁸ The toxic effects of lead occur at lower levels of exposure among lower-class children than among higher-class children.⁹ Although very high levels of exposure eventually overwhelm the status effect, what is remarkable is the consistency of the buffering effect of high status for entirely different forms of stress.

Example 2. What if education in the first two decades of life protected against dementia at the end of life? Until recently, this suggestion was treated as wild speculation,¹⁰ but sufficient evidence

FIGURE 2. Mental Development Index at 24 Months of Age by Umbilical Cord Lead Level and Social Class



Source: David Bellinger et al., "Low-Level Lead Exposure, Social Class, and Infant Development," *Neurotoxicology and Teratology* 10 (1993): fig. 1d.

has now accumulated for it to be taken seriously. Education has been found to be a protective factor both against dementia (in six different parts of the world)¹¹ and against a decline in mental function in late life.¹² This is not as farfetched as it may seem. After all, learning and memory depend upon connections between nerve cells in various parts of the brain. Dementia involves the loss and disorganization of these connections. Education enriches the network of interconnections, creating reserve capacity which can compensate for the inevitable losses that occur with aging; it seems probable then that education can postpone the onset of dementia. Moreover, once the connections are made, a "use it or lose it" principle may apply to our learning and memory capacity.

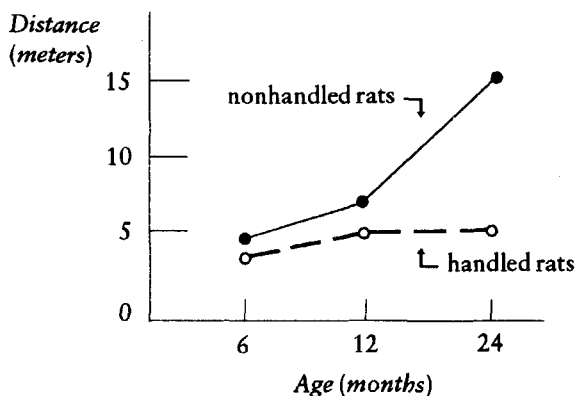
The risks of dementia and a decline in mental functioning seem to be greater among illiterate people than among primary school graduates, and greater among primary school graduates than among high school graduates. At first glance this may appear to follow a

“dose-response” pattern: the greater the education, the richer the nerve cell interconnections. But it could support equally well the notion of a critical period in brain development, around the age at which individuals are taught to read, when lifelong protective interconnections in the brain are most efficiently made. This principle of critical periods of development is well established in primate species. For instance, cats who are not allowed to see out of one eye for the first several weeks of life will never be able to see out of that eye because the regions of the brain which perceive visual images are not stimulated at the time when they are able to learn their function.¹³

Example 3. This “critical periods” principle may also have some significance in terms of defining the causes of dementia. An experiment involving the handling of rats has provided valuable information in this area. Human handling of baby rats during the first three weeks of life has been shown to improve the rats’ stress response throughout life.¹⁴ After being handled, these rats produce lower quantities of stress hormones, which are damaging to the learning and memory centers of the brain, than do “nonhandled” rats. The learning and memory centers in “handled” rats, therefore, do not age as quickly as in nonhandled rats. By late life, “handled” rats continue to learn new tasks at the same rate as middle-aged rats, while “nonhandled” rats show marked age-related declines in new learning.

Figure 3 shows the results of a learning experiment in which rats had to find a submerged platform in a pool of opaque water, relying entirely upon cues from the surrounding room. It took nonhandled rats longer to accomplish this task as they grew older; no slowing occurred as the handled rats aged. What is most significant, however, is that the handling phenomenon could not be induced if handling started later than the first few weeks of life.¹⁵ In other words, there was a critical period early in the rats’ lives when they learned to cope with a stressful stimulus (human handling) that gave them a lifelong advantage. This lifelong advantage was not merely in the area of stress response, but crossed over into the domains of learning and memory.

FIGURE 3. Mean Swimming Distance to the Platform



Source: Robert M. Sapolsky, *Stress, the Aging Brain, and the Mechanisms of Neuron Death* (Cambridge, Mass.: The MIT Press, 1992), fig. 12.8.

MODELS OF THE LIFELONG IMPACT OF CHILDHOOD EXPERIENCES

The preceding examples hint at two different, and sometimes conflicting, explanatory models of the impact of childhood experiences in later life: the “latency” model and the “pathways” model.

The Latency Model

The latency model is based upon research findings which reflect the critical period principle. For instance, the risk of death from heart disease in the fifth decade of life is strongly associated with the size of an individual’s placenta at birth and weight gain during the first year of life.¹⁶ Similarly, certain early childhood stimulation programs have been effective in improving the life trajectories of disadvantaged children.¹⁷ What this model emphasizes is that discrete events, which tend to occur very early in life, will have a strong *independent* effect in later life. The common theme, therefore, is that in early life the right things must happen.

The critical periods principle could easily lead to a vaccination approach to policy. We could reason that the more important the critical periods are, the more important the latent effects become.

Ultimately, the more important the latent effects are, the more important the targeted strategies for child stimulation and social support become, at least during various critical periods in the development of the human brain. In other words, socioeconomic influences on health, well-being, and competence throughout the life cycle could be addressed through a series of "magic bullet" type social, emotional, and/or educational strategies: highly focused in the developmental stage when they are applied and very specific in content. Deeper challenges to the nature of society could be blissfully ignored.

The Pathways Model

The pathways model is closely associated with the findings from long-term follow-up studies of newborns, adolescents, working populations, and the elderly. When these studies are compiled, the pattern that emerges highlights the lifelong impact of status on health, well-being, and competence. In essence, this model emphasizes the *cumulative* effects of life events and the ongoing importance of the conditions of life. For example, status differences at birth are associated with different levels of stability and security in early childhood, which are, in turn, associated with different levels of readiness for schooling.¹⁸ Lack of readiness for schooling leads to an increased risk of behavioral problems in school and ultimately to failure.¹⁹ Behavioral problems and failure in school lead to low levels of well-being in early adulthood.²⁰

In addition, parental status helps to determine the community where a child grows up, which, by the early school years, starts to influence the child's life chances through the social networks, community values, and opportunities which present themselves.²¹ By early adulthood individuals start to define their own status and differences begin to emerge: those who are doing better report higher levels of well-being. As adulthood unfolds, lower status individuals tend to end up in jobs which make relatively high demands on them, but which offer low levels of control concerning the pace and character of the work.²² By the fifth decade of life, people employed in these high demand jobs begin to develop high rates of disability and absenteeism,²³ eventually leading to early deaths.²⁴ This general pattern, which is more pronounced among

those who are socially isolated, persists into the eighth decade of life.²⁵

LATENCY, PATHWAYS, AND WORLDVIEWS

At first glance, it would appear commonsensical to view the latency and pathways models as complementary. After all, there is no reason to suppose that latent factors only act independently; any early life event which could exert a latent effect could also be the first step along a lifelong pathway which might have implications for health, well-being, or competence in the future. Conversely, any early childhood intervention designed to improve health and well-being in the long run will occur within a specific context which will provide a mixture of opportunities or barriers. Indeed, the closer the correlation between early life events and subsequent lifelong pathways, the more difficult the methodological problem of estimating the separate contribution of each.

Consider, for example, the findings of the longitudinal follow-up study conducted on subjects from the Perry Preschool Study²⁶ (shown in *Table 1*) when they reached age nineteen. This study is significant because it was based upon one of the most comprehensive of the early intervention programs for children in American inner cities and because it was evaluated by comparing outcomes among children who were randomly assigned to "preschool" and "no preschool" groups. The data can be interpreted in two ways. One interpretation would emphasize the remarkable improvements experienced by the preschool group in relation to the controls: higher rates of school graduation, higher rates of college attendance, fewer arrests, fewer teenage pregnancies, higher rates of employment, and a lower proportion on welfare. The alternate viewpoint would emphasize the fact that most of the achievements of the preschool group, although impressive when compared with the controls, would not nearly match those of middle-class children who had greater opportunities but no special preschool experience. After all, 31 percent were arrested or detained at least once, 18 percent were on welfare at the time of the interview, and teenage pregnancy was still rampant.

TABLE 1. Major Findings at Age Nineteen in the Perry Preschool Study

Category	Number Responding	Preschool Group	No Preschool Group	p
High school diploma/GED	121	67 %	49 %	0.034
College or vocational training	121	38	21	0.029
Functional competence (average or above)	109	61	38	0.051
Ever classified as mentally retarded	112	15	35	0.014
Years in special education	112	16	28	0.039
Ever detained or arrested	121	31	51	0.021
Females only: teen pregnancies, per 100	49	64*	117*	0.076
Employed at age 19	121	50	32	0.031
Receiving welfare at time of interview	121	18	32	0.044

Total n=123

*These numbers should be interpreted as 0.64 and 1.17 teen pregnancies per person.

Source: David P. Weikart, "Early Childhood Education and Primary Prevention," *Prevention in Human Services* 6 (1989): table 1.

On one level, this example illustrates the complementary nature of the power of latency, as demonstrated by the effectiveness of the targeted "early life" intervention, and the power of the pathway, as demonstrated by the strict limitations to success which existed in the specific context of the study community. Yet, despite the conceptual complementarity, these two interpretations are in conflict with one another. The latency model leads to a vaccination approach, which serves as intellectual cover for those who want to avoid confronting questions about the nature of society. Those who support the pathways model, on the other hand, find that it is a convenient tool for prying open social policy questions and pursuing a broad agenda of social reform. This paradox illustrates Sylvia N. Tesh's concept of "hidden arguments," namely that "firmly but often unconsciously held answers to . . . questions [in this case, how human potential expresses itself and the nature of the obligations which members of society have to each other] guide scientists, policy makers, and ordinary citizens alike to different constellations of facts about the causes of disease and, hence, to different preferences for prevention policy."²⁷

The hidden arguments overlay a complex issue and distort it by framing the policy options in an "either/or" fashion. They reveal themselves when people show defensive responses to the notion that the human brain, like that of other primates, may have critical developmental periods. After all, who wants to embrace a model which says that "if they are not okay by age three, scrap them!" Hidden arguments are also at work among those who go to absurd lengths to deny the significance of status differences in health, well-being, and competence. Meanwhile, the orthodoxies survive, the arguments continue, and the problems remain unaddressed.

The worst possible outcome would be to reduce the argument to the notion that a cradle-to-grave social contract is the policy corollary of the pathways model and a one-shot early childhood social investment strategy is the policy corollary of the latency model. What is needed instead is a pragmatic approach which draws upon the strengths of each model.

Consider the issue of critical periods in human development. Humans, like other primate species, are predisposed to develop certain forms of competence at well-defined times in the life cycle. The most significant of these are aspects of cognitive and emotional functioning. In particular, certain stimuli are effective in promoting cognitive or emotional functioning at well-defined times within the first few years of life, but are ineffective either beforehand or afterwards. For humans, however, critical periods may be less critical than they are for other primates. There is consistent evidence, relating to a variety of areas of function, that if appropriate stimulation is missed at a specific time in early childhood, the function can be developed through other forms of stimulation later in life.²⁸ It may just be harder to do. For example, middle-aged adults can learn new languages as efficiently as young people if they are given the time and financial support to immerse themselves in the subject.²⁹ For this reason, as regards human development, the label "sensitive periods" is more appropriate than "critical periods." This distinction provides a biological basis for the continued relevance of both latent effects and pathways.

We need an approach to policy which builds towards a new working consensus. It should stress the importance of the life cycle. It should be action-oriented and grounded in social experimentation. And it should be receptive to the notion that there may be

small answers to big questions. The latter may be the most important characteristic of all. Many of the most promising social interventions are rather modest in character: the value of positive mentors and good recreation programs for the young, a few opportunities at a "second chance" to make a successful transition from childhood to adulthood, and a strong social support network to buffer the stresses of middle and late life.

ENDNOTES

¹Canadian Council on Social Development, *Newsletter of the Centre for International Statistics on Economic and Social Welfare for Families and Children* 1 (1992).

²Donna S. Lero, Alan R. Pence, Margot Shields, Lois M. Brockman, and Hillel Goelman, *Introductory Report: Canadian National Child Care Study* (Ottawa: Statistics Canada, 1992).

³Donna S. Lero, Hillel Goelman, Alan R. Pence, Lois M. Brockman, and Sandra Nuttall, *Parental Work Patterns and Child Care Needs* (Ottawa: Statistics Canada, 1992).

⁴Anton E. Kunst and Johan P. Mackenbach, *An International Comparison of Socio-Economic Inequalities in Mortality* (Rotterdam: Erasmus University, 1992); Anton E. Kunst, J. J. M. Guerts, and J. Berg, *International Variation in Socio-Economic Inequalities in Self-Reported Health* (The Hague: Netherlands Central Bureau of Statistics, 1992); and Clyde Hertzman, "Where are the Differences Which Make a Difference?," *CIAR Population Health Working Paper* 8 (1990).

⁵In this paper, I replace "gradient" with the more colloquial "doing better" or "doing worse." I use "status" to refer to any measure of socioeconomic status or social class since the research shows that they are effectively interchangeable.

⁶Bill Moyers, *Healing and the Mind* (New York: Doubleday, 1993); Robert M. Sapolsky, *Stress, the Aging Brain, and the Mechanisms of Neuron Death* (Cambridge, Mass.: The MIT Press, 1992).

⁷Emmy E. Werner, "Children of the Garden Island," *Scientific American* (April 1989): 106-111; Emmy E. Werner and Ruth S. Smith, *Overcoming the Odds: High Risk Children from Birth to Adulthood* (Ithaca, N.Y.: Cornell University Press, 1992).

⁸Herbert L. Needleman and David Bellinger, "The Health Effects of Low Level Exposure to Lead," *Annual Review of Public Health* 12 (1991): 111-40.

⁹David Bellinger, Alan Leviton, Christine Wateraux, Herbert Needleman, and Michael Rabinowitz, "Low-Level Lead Exposure, Social Class, and Infant Development," *Neurotoxicology and Teratology* 10 (1993): 497-503.

¹⁰As recently as three years ago, reviewers for at least one relevant journal were treating such ideas as amusing fiction.

- ¹¹Robert Katzman, "Education and the Prevalence of Dementia and Alzheimer's Disease," *Neurology* 43 (1993): 13-20.
- ¹²Ibid. Denis A. Evans, Laurel Beckett, Marilyn Albert, Liesi Hebert, Paul Scherr, Harris Funkenstein, and James Taylor, "Level of Education and Change in Cognitive Function in a Community Population of Older Persons," *Annals of Epidemiology* 3 (1993): 71-77.
- ¹³Ronald E. Kalil, "Synapse Formation in the Developing Brain," *Scientific American* (December 1989): 76-85.
- ¹⁴Sapolsky, *Stress, the Aging Brain, and the Mechanisms of Neuron Death*. M. Meaney, D. Aitken, S. Bhatnager, C. van Berkel, and Robert Sapolsky, "Effect of Neonatal Handling on Age-Related Impairments Associated with the Hippocampus," *Science* 239 (1988): 766.
- ¹⁵Sapolsky, *Stress, the Aging Brain, and the Mechanisms of Neuron Death*.
- ¹⁶David Barker and C. Osmond, "Infant Mortality, Childhood Nutrition, and Ischaemic Heart Disease in England and Wales," *The Lancet*, 10 May 1986, 1077-81; David Barker, C. Osmond, Jean Golding, Di Kuh, and Michael Wadsworth, "Growth in Utero, Blood Pressure in Childhood and Adult Life, and Mortality from Cardiovascular Disease," *British Medical Journal* 298 (1989): 564-67; David Barker, C. Osmond, P. Winter, B. Margetts, and S. Simmonds, "Weight in Infancy and Death from Ischaemic Heart Disease," *The Lancet*, 9 September 1989, 577-80; David Barker, A. Bull, C. Osmond, and S. Simmonds, "Fetal and Placental Size and Risk of Hypertension in Adult Life," *British Medical Journal* 301 (1990): 259-62; David Barker, "The Intrauterine Environment and Adult Cardiovascular Disease," in Ciba Foundation Symposium 156, *The Childhood Environment and Adult Disease* (Chichester: Wiley, 1991), 3-16; David Barker and C. Martyn, "The Maternal and Fetal Origins of Cardiovascular Disease," *Journal of Epidemiology & Community Health* 46 (1992): 8-11; David Barker, K. Godfrey, C. Osmond, and A. Bull, "The Relation of Fetal Length, Ponderal Index and Head Circumference to Blood Pressure and the Risk of Hypertension in Adult Life," *Pediatric and Perinatal Epidemiology* 6 (1992): 35-44; David Barker, T. Meade, C. Fall, A. Lee, C. Osmond, K. Phipps, and Y. Stirling, "Relation of Fetal and Infant Growth to Plasma Fibrinogen and Factor VII Concentrations in Adult Life," *British Medical Journal* 304 (1992): 148-52; and David Barker, C. Osmond, S. Simmonds, and G. Wield, "The Relation of Small Head Circumference and Thinness at Birth to Death from Cardiovascular Disease in Adult Life," *British Medical Journal* 306 (1993): 422-26.
- ¹⁷David P. Weikart, "Early Childhood Education and Primary Prevention," *Prevention in Human Services* 6 (1989): 285-306; Francis H. Palmer, "Long-Term Gains from Early Intervention: Findings from Longitudinal Studies," in Edward Zigler and Jeanette Valentine, eds., *Project Head Start: A Legacy of the War on Poverty* (New York: The Free Press, 1979); and Sandra L. Martin, Craig T. Ramey, and Sharon Ramey, "The Prevention of Intellectual Impairment in Children of Impoverished Families: Findings of a Randomized Trial of Educational Day Care," *American Journal of Public Health* 80 (1990): 844-47.
- ¹⁸Robbie Case and Sharon Griffin, "Rightstart: An Early Intervention Program for Insuring that Children's First Formal Learning of Arithmetic is Grounded in their

Intuitive Knowledge of Numbers," Report to the James S. McDonnell Foundation, 1991.

¹⁹Lea Pulkkinen and Richard E. Tremblay, "Patterns of Boys' Social Adjustment in Two Cultures and at Different Ages: A Longitudinal Perspective," *International Journal of Behavioural Development* 15 (1992): 527-53; and Chris Power, Orly Manor, and John Fox, *Health and Class: The Early Years* (London: Chapman & Hall, 1991).

²⁰Power et al., *Health and Class: The Early Years*.

²¹Mary Haan, George A. Kaplan, and Terry Camacho, "Poverty and Health: Prospective Evidence from the Alameda County Study," *American Journal of Epidemiology* 125 (1987): 989-97.

²²Robert Karasek and Tores Theorell, *Healthy Work: Stress, Productivity, and the Reconstruction of Working Life* (New York: Basic Books, 1990).

²³Michael Marmot, "Explaining Socioeconomic Differences in Sickness Absence: The Whitehall II Study" (Toronto: Canadian Institute for Advanced Research, 1993); Michael Marmot, George Davey-Smith, Stephen Stansfeld, Chandra Patel, Fiona North, Jenny Head, Ian White, Eric Brunner, and Amanda Feeney, "Health Inequalities among British Civil Servants: The Whitehall II Study," *The Lancet*, 8 June 1991, 1387-393.

²⁴Michael Marmot, M. Kogevinas, and M. Elston, "Social/Economic Status and Disease," *Annual Review of Public Health* 8 (1987): 111-35.

²⁵Michael Wolfson, Geoff Rowe, Jane Gentleman, and Monica Tomiak, "Career Earnings and Death: A Longitudinal Analysis of Older Canadian Men," *Population Health Working Paper No. 12* (Toronto: Canadian Institute for Advanced Research, 1991).

²⁶Weikart, "Early Childhood Education and Primary Prevention."

²⁷Sylvia N. Tesh, *Hidden Arguments* (New Brunswick, N.J.: Rutgers University Press, 1988), 1.

²⁸Marc H. Bornstein, ed., *Sensitive Periods in Development: Interdisciplinary Perspectives* (New Jersey: Lawrence Erlbaum Associates, 1987).

²⁹Ibid.

Toward a System of Health Statistics

IT IS A COMMONPLACE THAT WE LIVE in the “information age.” Yet, in the area of health we have inherited a hodgepodge of data collection systems that are seriously limited with respect to current thinking. This is particularly true for the range of determinants of health discussed elsewhere in this volume. There is more to a health information system than the sophisticated workstations, networks, and graphical user interfaces, which, it is proposed, will send magnetic resonance images of a patient’s head around the continent much as one would a phone call. What is needed, and what this paper aims to provide, is a broader vision of health information—a system of health statistics—in line with emerging views on the determinants of health.

As a preliminary, one semantic point. The words *data*, *statistics*, and *information* are used frequently. They are not interchangeable. Data is best thought of as the raw material for making statistics and information. In turn, statistics are typically quantitative, and thus are a subset of information.

CONVENTIONAL HEALTH INFORMATION

Health data, particularly vital statistics, such as cause of death, are among the oldest types of organized systems of data collection in Canada and in many other countries. Canada has one of the best sets of health data in the world, but we can do much better. Following is a brief, critical review of a few major areas of current health information activity.

Michael C. Wolfson is a senior official at Statistics Canada and a Fellow of The Canadian Institute for Advanced Research.

The first, and perhaps the most common, kind of health information is broad macroindicators relating to the health of large population groups. Two of the most widely used macroindicators are life expectancy and health-care spending (as a percentage of the gross domestic product [GDP]). These indicators form the basis for virtually all international comparisons. This is not because better indicators have not been proposed, but because the historical importance of the underlying National Accounts, population census, and death registration data make them readily available. These are, however, curious indicators of population health. Life expectancy (essentially average age at death) takes no account of the health status of the living. Health-care spending is a measure of resource inputs, where if we were really healthy as a population, less would be better.

At a more micro level, hospitals are major creators of health-related data. As part of their routine operations, they keep track of the number of patients filling their beds, the number and type of surgical procedures they perform, and, of course, their finances. Moreover, armies of medical records clerks are employed to produce computerized abstracts of the files of each patient who passes through the hospital's doors. These data have become important as part of the general pressure to control costs. How much funding a hospital receives depends increasingly on how severely ill its patients are, and whether hospitals are adhering to guidelines for appropriate treatment. However, routinely collected data on patients rarely include the kinds of socioeconomic or psychosocial factors known to be important determinants of health. Nor are patients generally followed after they leave the hospital. Thus, hospital-based data tend to have a short and narrow perspective.

There is an industry of medical researchers, substantially funded by governments and pharmaceutical firms, who run clinical trials. They create and collect detailed information on the effects of myriad interventions, based mainly on the paradigm of randomized experimentation. Most of these intervention experiments are very narrowly focused, concentrating on specialized end points in small, nonrepresentative populations (i.e., does drug X affect blood fraction Y by at least P percent in people suffering from disease Z?). Randomization is designed to relieve the experimenter of the need to understand the complex determinants of health. But because the randomization research paradigm is unable to handle a more com-

plex *system* view that includes, for example, many contending causal pathways and diverse feedback processes, this sort of research is drawn toward "magic bullet" interventions.

The most powerful kind of data is both broad enough to be representative of the entire population and detailed enough to reflect the heterogeneity of individuals. Hospitals produce countless tidbits of data (individual microrecords) each year from records of discharges, physician visits, and from certificates of births and deaths, but each record tends not to contain much breadth of information and is stored as an isolated event. Many countries have periodic population health surveys to track, for example, self-reported chronic disease burdens and health-related behaviors like smoking. However, there is almost no connection between the clinical data from hospitals and the self-reported vernacular descriptions of health problems from health surveys.

EMERGING HEALTH INFORMATION

One of the central policy questions in the health area at present is how to allocate resources in the most cost-effective manner. Within a hospital, the question may be framed as the relative costs and benefits of the allocation of beds and surgical suite time between two different groups of high tech health-care providers, such as cardiology and orthopedics. At the cabinet level in provincial governments, the important trade-offs are not so much between specific kinds of health care as between funding health care generally and other approaches to improving the population's health status, such as support for early childhood development.

Another area of emerging health information needs is comparative reporting. One example is variations between geographic areas in surgical procedure rates. The obvious question that has been generated by these kinds of data is why residents in some neighborhoods have a greater chance of having a particular organ removed than do residents of similar neighborhoods. The answer is still contested, but none of the obvious hypotheses about one neighborhood being chronically sicker or having less access to health care than another have much support in the data.

Clearly, more information is needed than is currently provided by isolated visit or encounter data. To support even the most basic

kinds of analysis, data on each individual's health service visits or encounters must be strung together to form person-oriented data. However, connecting various bits of such data together will be meaningless without common concepts and definitions, and without a broad population perspective.

At the health macroindicator level, there is an expanding consensus that life expectancy data are simply not enough. There are major concerns about chronic disease burdens and the growing population of frail elderly. There is also debate regarding James Fries' "compression of morbidity" hypothesis—essentially whether the additional years of our increasing longevity are predominantly healthy or sickly.

There is also a need for more reliable observation of health inequalities. Evidence increasingly shows that the correlation of health and wealth is not only short term. Health inequalities are associated with long-term effects and therefore cannot simply be observed at a point in time. For example, longitudinal data on individual workers show that those with higher incomes in middle age enjoy greater longevity, living well past retirement.¹ Not dissimilarly, children born to parents with lower social status have more frequent behavior problems and poorer health in their early adult years.²

Such long-term and significant latent effects in turn raise fundamental questions about causal pathways. What are the roles, for example, of homeostasis, resilience, coping skills, and Aaron Antonovsky's "sense of coherence" as basic explanations for continuing good health?³ How can we balance our current preoccupation with disease with information on positive health? Perhaps we should focus on health as a dynamic concept based on repeated longitudinal observations. Then the question would not be "why is Jones sick today?" but "why is Smith always healthy?"—or more generally, what are the determinants of "chronic good health"?

The significant and long lasting correlations between socioeconomic status and health problems imply that we must view diseases as processes of long duration with a range of determinants (multifactorial, interacting, dynamic) going well beyond medicine and biochemistry. An intriguing example is psycho-neuro-immunology (PNI), which, from a lay perspective, gives credence to the premise of strong but subtle connections between mind and body. Major

advances in understanding may result from connecting this emerging knowledge on biochemical pathways to population-based surveys. For example, a population survey could ask people for subjective perceptions of their health and the psychosocial aspects of their home and workplace milieus, collect blood samples and analyze biochemical markers of chronic stress, and then follow respondents over a number of years. Something along these lines is being discussed in connection with the new Canadian National Population Health Survey. Such new research needs will require changes in our basic health research institutions and a strengthened role for population-based health information.

DIAGNOSES OF DATA: CONCEPTUAL PROBLEMS

There are two fundamental problems with the health data that are currently being collected. The first of these is a pervasive *imbalance*. We know far more in Canada about how much the health-care system costs than we do about how healthy the population is. We have regular data on health-care spending, life expectancy, and incidences of defined disease. But do we know if the living population is healthier now than it was a few years ago? Do we even have sufficient data on health outcomes to assess systematically which medical or surgical procedures actually improve health?

The preponderance of data currently collected on health care derives from the "medical model" of health and is out of balance with our knowledge of the much broader range of the determinants of health. For example, there has been a great deal of epidemiological research on the relationship between elevated cholesterol levels and coronary heart disease but far less on the mechanisms by which low cholesterol levels are associated with elevated all-cause mortality.⁴

One important corollary of the paucity of health outcome data relates to health goals. Ideally, we should define our provincial, state, or national goals first, and then develop the necessary information systems to monitor progress toward those goals (e.g., not only adding years to life, but also adding life to years). Unfortunately, I think we sometimes fail to express or even to think about such basic health goals simply because we do not have the required data.

The second problem is a bit more abstract, but equally fundamental. We have sets of health statistics, not a *system*, and our

health statistics suffer from incoherence. An indication of this incoherence can be gained by perusing any reasonable compendium of health statistics, and for convenience I could pick on "Health United States"⁵ or Statistics Canada's set of health indicators.⁶ In both cases, the numbers are a hodgepodge. Nothing adds up; there is no integrating framework, as exists for the System of National Accounts. One page of "Health United States" shows bed days, another shows infant mortality, another displays numbers of health professionals, and yet another dollars spent. The only thing holding these statistics together is a general pertinence to health, and the binding of the publication itself!

FURTHER DESIDERATA

Imbalance and lack of coherence in our health information system must be countered by a broadly-based, outcome-oriented population perspective, common concepts and definitions, and an appropriate mathematical structure.

Health and health-related activities should be measured in their own most natural units. It is unnatural to measure health in dollars. The central concept in a system of health statistics should be population health status. The natural unit of measure for health would then be something like a "health status adjusted" life year. Of course, such a central concept should be complemented by a variety of other measures, such as numbers of positive health behaviors (e.g., days with physical activity), numbers of health-care encounters of various kinds, economic costs, and diseases.

A second major desideratum is the need to compile not only aggregate health information but individual-level health information. The health information system should have explicit microdata foundations. Individuals are so heterogeneous in their genetic predispositions, socioeconomic status, and health histories that information for representative samples is critical to analyses of health inequalities, and to understanding more broadly.

This may seem self-evident, but in the National Accounts (the premier statistical system in Canada and most countries) as well as in demographic estimates and projections, the smallest level of disaggregation is a *group* (e.g., an industrial sector or an age range). The semi-aggregate character of these statistical systems is under-

standable, given that they both predate the current revolution in computing. However, it is an unfortunate limitation that should and can be avoided in the development of any new statistical systems, particularly those for population health.

One final desideratum is taking full advantage of modern informatics. A new system of statistics should be much more than a series of print publications. Its full incarnation should be an electronic database, combined with tailored retrieval and analytical software. The ideal system should offer the user an interface whereby she or he can ask questions and get answers (subject always to constraints imposed by considerations of privacy and confidentiality).

THEORIES AND CONCEPTS OF HEALTH

In thinking of a new or substantially redesigned health information system, we must ask what is important to measure, and how it ought to be measured. Answers to these fundamental questions necessarily derive from a theoretical or conceptual framework. For example, theories in astrophysics explicitly guide the development and construction of specific kinds of radio telescopes. In the case of human health, we are at a point where there is enough accumulated evidence on the broad range of the determinants of health to signal the need for major changes in theoretical and conceptual perspective and consequential changes in approaches to observation and measurement.

We know, for example, that social support and incomes are connected to longevity. We also know that there are specific "windows" in early childhood when key opportunities for physiological development (e.g., vision) are present. However, we have no grand theories that knit these diverse empirical findings together, or that have driven observational developments in a way similar to those in fields such as astrophysics. A system of health statistics should be based upon a conceptual framework that is broad enough to encompass all these kinds of observations.

THE HEALTH INFORMATION TEMPLATE

There is, of course, a rich and significant history of conceptual frameworks for health.⁷ Given the ferment of new ideas in Canada,

represented in part by the work of members of the Population Health Program of The Canadian Institute for Advanced Research, and a broadening appreciation of the malaise in our current health information system, a National Task Force on Health Information was organized several years ago under the direction of the previous Chief Statistician of Canada.⁸ Very early in the work of the Task Force it was agreed that a "road map" or conceptual framework setting out the territory to be covered was essential. Peter Glynn and I embarked on this task, convening an ad hoc working group under the aegis of the Task Force. The result is a graphical software application called the Health Information Template.⁹

The Template has served a variety of purposes. One is simply as a pedagogical device to assist the Task Force in exploring and hopefully reaching consensus on what should be considered part of the field of health information. A second purpose of the Template is to provide the beginnings of a classification structure for the myriad of data, statistics, and information under discussion. Finally, the Template provides the core of a conceptual framework, which can be used not only as a way of indexing various types of information, but also as a structure for indicating how different types of information are connected to one another.

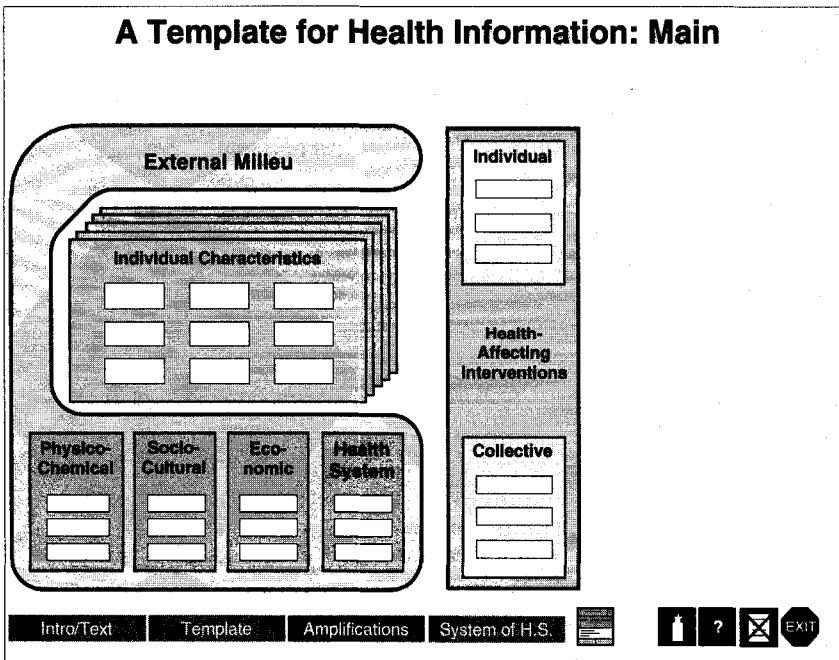
THE MAIN TEMPLATE IMAGE

Figure 1 shows a monochrome version of the main template image, the "road map" used by the National Task Force on Health Information. This template divides the health field into three broad domains. At the center of the image is a population of individuals and its characteristics. The files going back into the page indicate that the data for the population must be derived from individual-level microdata.

Surrounding, but not completely enveloping the population of individuals (both visually and metaphorically), is the "External Milieu." In other words, the many milieux or environments within which individuals live are essential to the domain of health information. The "External Milieu" is classified into four kinds of environments. The most popular association with the word "environment" in the health context is that of physicochemical exposures. However, for people in the advanced industrial societies, exposures to physi-

cochemical hazards like lead and smog do not play as important a role as one might think. Far more important is the sociocultural environment that includes such things as our relationships with family and friends, television images of violence, and peer pressures on teenagers. There is also clear evidence that the economic environment of income and work has a major influence on health. Simply put, individuals who earn higher incomes live longer and are healthier. In fact, the economic environment has as great an effect on health as do smoking and high levels of cholesterol. The fourth kind of environment which comprises the "External Milieu" is the health system itself, specifically those aspects inherited from the past, like existing stocks of equipment and hospital buildings.

FIGURE 1.



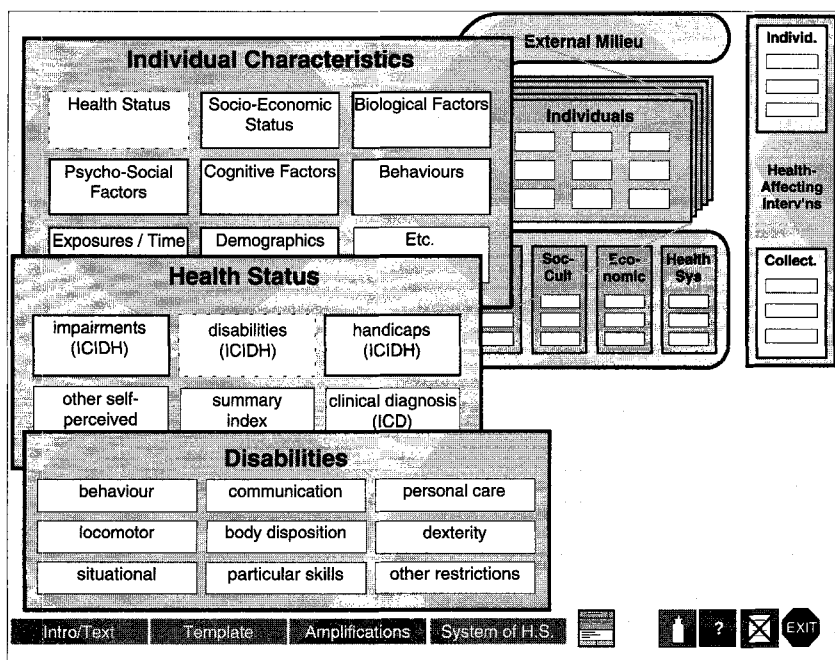
The bulk of society's health-related activity is shown in the vertical bar to the right of the image, implying that while we are surrounded by the "External Milieu," our health is not entirely determined by environmental exposures. Health is influenced by a

variety of health-care services, and by a broader range of "Health-Affecting Interventions." These interventions can be logically divided into two groups. The first, "individual" interventions, are typically one-on-one encounters with providers of various health-care services. The second, "collective" health-affecting interventions, are government programs and regulations that act on us collectively, though indirectly (and often inadvertently), through the "External Milieu." Examples include *regulations* of food quality, tobacco advertising, pharmaceuticals, and particulate emissions, as well as *expenditures* on traffic safety and head start programs.

TEMPLATE: MAIN IMAGE (EXPLODED)

The white boxes shown within each of the domains and subdomains of the main template image represent repositories for various categories of information which are based on finer classification structures. Figure 2 shows the result of "opening" one of the various domains and subdomains of the template.

FIGURE 2.



It should be emphasized that the boxes at various levels represent labels in a hierarchical classification structure, and not compartments. During the consultation process of the Health Information Task Force, we found that some groups, for example those in health promotion where individuals' backgrounds were primarily qualitative, had difficulty with this image. Since many health promotion advocates support a more holistic approach to health, they feared that the Template was trying to compartmentalize various aspects of health and health information, and thereby prevent integration. Quite the contrary. The notion of explicit microdata foundations, the idea that information should be collected and maintained at the individual level, means that integrated analysis is not only possible, but that its potential is maximized. The intent of the Template is very much to provide a holistic view.

INFERENTIAL APPARATUS

The two images of the template (*Figures 1 and 2*) are static. But clearly, health is a dynamic process. The processes of expressing genetic predispositions and reacting to risk factor exposures and disease progression, for example, all evolve over time. Moreover, underlying biomedical phenomena such as the transformation of a precancerous cell into a tumor often exhibit very long latencies. Salient health information should include descriptions of these kinds of dynamic processes—and not just in prose form. Relevant information includes statistical descriptions of patterns of onset and rates of progression for various kinds of health problems, their precursors, and their sequelae.

Given such dynamic or process descriptions, a major use of this kind of information system is to generate projections. For example, astronomers (very successfully) use their understanding of orbital dynamics to project the time and location of solar eclipses. Economists (not always as successfully) use macroeconomic models to assess the likely impacts of one or another deficit reduction strategy. An example of such quantitative projection or inference regarding likely effects in the health area is the MRFIT (Multiple Risk Factor Intervention Trial). In this \$100 million study, risks of developing coronary heart disease were taken from the published epidemiological literature. The observed risk patterns were then used in the

design of MRFIT to predict (not very successfully) the impact of interventions for lowering cholesterol and cessation of smoking in a large US population.

These kinds of projections imply a tight coupling between an information system and some sort of "inferential apparatus." Typically, an "inferential apparatus" is a computer simulation model. Functionally, it is a systematic, rigorous, and replicable method for drawing out the implications of data under hypothesized scenarios. Implicitly, if not explicitly, virtually all data are actually collected in order to support modeling. This may seem a sweeping statement, but any collection of data presupposes some model or conceptual framework into which the user will place those data.

Physicists and economists have a long history of "grand theories" (e.g., general relativity, Keynes' general theory) that underlie both extensive programs of data collection and large scale modeling efforts. However, no such grand theory exists in the area of human health; and the conceptual frameworks that do exist are too general.

We thus face a conundrum. On the one hand, it has just been argued in effect that all important information systems are tightly coupled to some sort of method for making projections and answering "what if" questions—where the method nowadays often takes the form of a computer simulation model which in turn represents a theory. On the other hand, we seem not to have any theory in health that is anywhere near as broad in scale and scope as the information spanned by the Health Information Template, particularly as shown in *Figures 1 and 2*.

To avoid presuming a general theory of the health of human populations, I shall refer instead to the notion of an inferential apparatus which transforms health-related data into projections. The Health Information Template includes a sketch of such an apparatus. In part, this provides the symbiote for the broad range of data collection implicit in the Template.

How might we conceive an inferential apparatus for the health of human populations? One approach is to borrow the core concept used in descriptions of dynamic systems, whether in astrophysics or evolutionary economics. This is the idea of "state variables" which evolve over time according to well-specified dynamics or "laws of motion." In the case of health information, the state variables can be any number of attributes of individuals in a population, such as

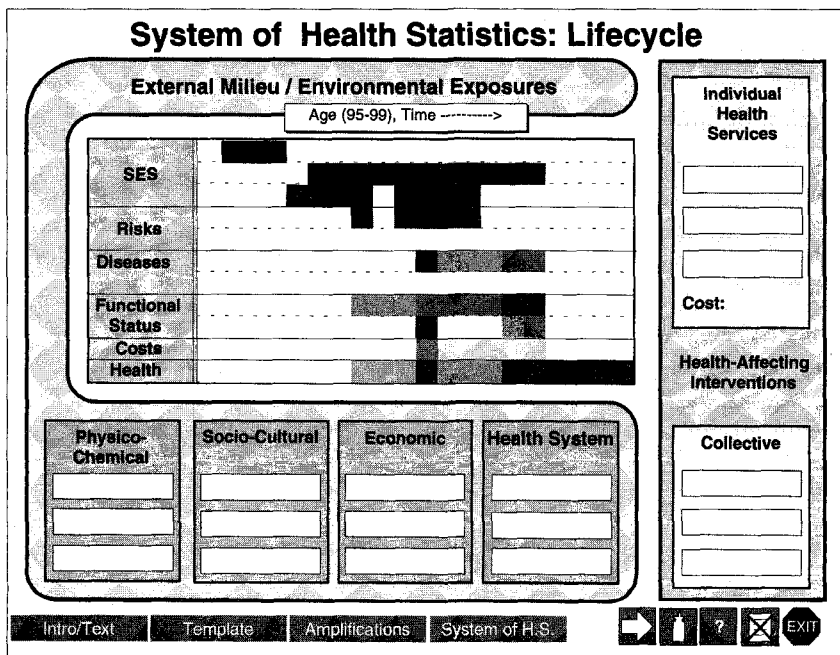
age, cholesterol level, occupation, stress level, and coronary heart disease status.

Of course, most of the dynamics of human health are not well understood, so it is a huge leap to assume the existence of well-defined laws of motion for the various attributes of human health. But that is not the point. The intent, rather, is to have a framework that can absorb and utilize such information as it is generated and, equally importantly, that can guide the development of new health-related data and information. A key implication is that the core of our health information system should be based on longitudinal microdata on individuals, giving a life-cycle perspective.

TEMPLATE: SYSTEM OF HEALTH STATISTICS LIFE-CYCLE IMAGE

Figure 3 shows such an individual life-cycle perspective. The matrix of cubbyholes in the central part of the image is like a formally structured biography—each row corresponds to a group of state variables, and each column refers to an age interval along the horizontal axis. This image is also like the record layout of a computerized longitudinal microdata set.

FIGURE 3.



The main groups of variables describing an individual's stylized biography are: 1) socioeconomic status (SES), including whether the individual is in school, married, and/or in the paid labor force; 2) various risk factors like smoking, genetic predispositions, and availability of social support; 3) clinically defined diseases such as chronic heart disease, lung cancer, arthritis, and Alzheimer's disease; 4) vernacular health problems such as not having sufficient lower limb mobility or cognitive capacity to get around without a wheel chair; 5) direct costs of health services (high use being indicated in the diagram by darker shading) as well as economic costs (i.e., foregone earnings); and 6) a summary health status value (i.e., a number between zero [black=dead] and one [white=fully healthy]) summarizing each individual's overall health status for the year.

The dynamics of the state variables in an individual's biography are not explicit in this image. All that is shown are the top levels of a classification structure. However, "laws of motion" like statistical descriptions of disease incidence and survival probabilities, as well as the quantitative contributions of various risk and protective factors, are fundamental to a complete system of health statistics. These kinds of statistical descriptions are being assembled as part of Statistics Canada's POpulation HEalth Model (POHEM), which is also an "inferential apparatus" as described earlier.¹⁰ POHEM is being applied in a number of areas of health analysis and statistics: extending the concept of life expectancy to produce indices of "healthy life expectancy" (HLE), estimating the burdens of chronic diseases like arthritis in terms of HLE, assessing the coherence of existing data in the case of lung cancer, estimating the population health impacts of interventions like cholesterol lowering, and exploring the quantitative importance of socioeconomic factors.

CONCLUSION

The Health Information Template provides a very ambitious image of a health information system, one that is coherent and balanced and is able to encompass a broad view of the determinants of health. I am hopeful that this image can be realized. As part of a broad set of health information initiatives, our Population Health Model is moving beyond prototype to support applied analysis; we have a new Canadian Institute for Health Information being born;

and Statistics Canada has new resources to support innovations in health information, including a new longitudinal National Population Health Survey. At the same time, the discussions reported in this volume which help deepen our understanding of the health of populations suggest exciting, new, and as yet unfunded health information challenges.

ENDNOTES

- ¹Michael C. Wolfson, Geoff Rowe, Jane Gentleman, and Monica Tomiak, "Career Earnings and Death—A Longitudinal Analysis of Older Canadian Men," *Journal of Gerontology* 48 (4) (1993): S167–S179.
- ²Chris Power, Orly Manor, and John Fox, *Health and Class: The Early Years* (London: Chapman and Hall, 1991).
- ³Aaron Antonovsky, "Can Attitudes Contribute to Health?," *ADVANCES, The Journal of Mind-Body Health* 8 (4) (Fall 1992): 33–49.
- ⁴John Frank, Dwayne M. Reed, John S. Grove, and Richard Benfarte, "Will Lowering Population Levels of Serum Cholesterol Affect Total Mortality?—Expectations from the Honolulu Heart Program," *Journal of Clinical Epidemiology* (1992).
- ⁵National Center for Health Statistics, "Health United States," DHHS Pub. No. (PHS) 86–1232, Washington, D.C., US Government Printing Office.
- ⁶Statistics Canada, "Health Indicators," Ottawa, Statistics Canada.
- ⁷An excellent historical review and new, broader conceptual framework based on the determinants of health has been developed by Evans and Stoddart. See Robert G. Evans and Gregory Stoddart, "Producing Health, Consuming Health Care," *Social Science and Medicine* 31 (12) (1990): 1347–363. See also Robert G. Evans, M. L. Barer, and Theodore R. Marmor, *Why are Some People Healthy and Others Not? The Determinants of the Health of Populations* (New York: Aldine de Gruyter, 1994).
- ⁸Martin Wilk, "Report of the National Task Force on Health Information" (Ottawa: Statistics Canada, 1991).
- ⁹Michael C. Wolfson, "A Template for Health Information," *World Health Organization Statistical Quarterly* (Geneva: WHO, March 1992). A gratis copy of the Template software is available.
- ¹⁰Michael C. Wolfson, "POHEM—A Framework for Understanding and Modeling the Health of Human Populations," US Census Bureau Annual Research Conference, Proceedings, 261–82.

It is almost "un-American" to look into the future and find disappointment and failure. Efforts at reform are supposed to produce successes that will improve life for the individual and the society. Nevertheless, while the health-care system is responsive to the most urgent needs of most of the population, there is no prospect whatever, no matter how much money we invest in it, for the system to serve as broad support for the host of ills that already beset modern man and that may well multiply in the years ahead. In consuming over 8 per cent of the gross national product, the health-care system is crowding the margins of family and public income, and we should consider carefully what we are likely to buy if we increase our expenditures to any significant degree.

Questions of investment aside, we must also consider whether we have, or are capable of designing, the social machinery needed to exercise tighter control over the uses to which health resources are put. The streets have not been made safe for pedestrians; the schools have not learned to teach many pupils how to read and write; and corruption permeates much of our public and business life. In light of these failings in the provision of services, a call for moderation in the setting of health goals and for realism in the assessment of power and decision do not represent a doctrine of despair and defeat so much as a plea for the role of reason in the search for a better society.

Eli Ginzberg

From "Health Services, Power Centers, and
Decision-Making Mechanisms"
Dædalus 106 (1) (Winter 1977)

Social Differentials in Health Within and Between Populations

INTRODUCTION

BETWEEN 1965 AND 1990, LIFE EXPECTANCY for both men and women improved by six years (from sixty-eight to seventy-four and from seventy-four to eighty, respectively) in the Organization for Economic Cooperation and Development (OECD) countries. There were even greater improvements in much of the rest of the world: eight years for men and nine years for women (from forty-one to forty-nine and from forty-three to fifty-two, respectively) in sub-Saharan Africa; and fourteen years for both men and women (from forty-six to sixty and from forty-four to fifty-eight, respectively) in India. It is a fair guess that improvements in life expectancy are attended by improvements in health status.

Why are things getting better? This question is of much interest to demographers but, judging by papers in the major medical journals, of little interest to doctors, with a few notable exceptions.¹ We all have ready answers to why life expectancy has improved in the poor countries of the world: improved sanitation, immunization, health education, medical care, food, education, and economic development. (These answers, though obvious, may not all be correct.) The situation is more complex. It is not so easy to produce an answer to why life expectancy has improved in the richer countries of the world.

If we believe that there is a link between wealth and health, one possible answer is that the rich countries have become richer. There are two problems with this rather simple approach. First, we do not

Michael G. Marmot is Professor of Epidemiology and Public Health and Head of the Department of Epidemiology and Public Health at University College London Medical School, England.

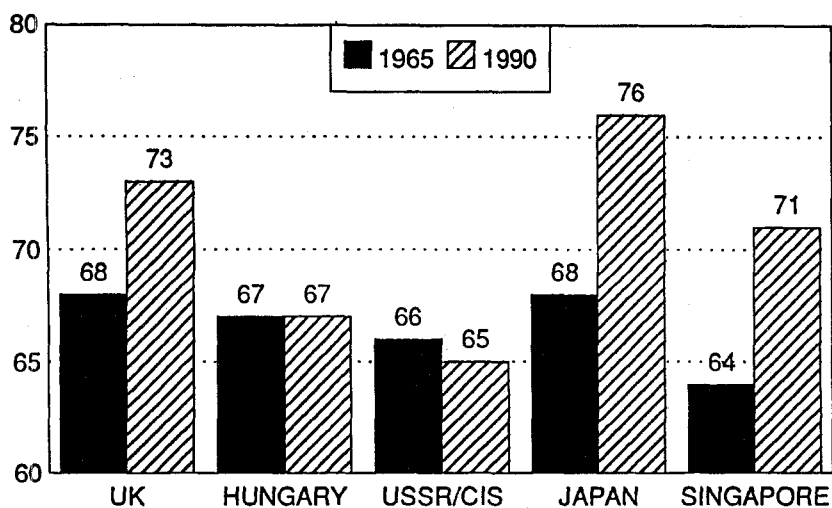
understand very well the reasons for the link between wealth and health. Second, as Richard Wilkinson has shown, there is a close relationship between Gross National Product per capita (GNPpc) and life expectancy up to a ceiling of about \$5,000. Beyond that, there is little relation between a country's GNP and the life expectancy of its inhabitants.² Wealth is important but it alone does not provide a complete explanation for the improvements in life expectancy.

Another puzzling aspect of the improvement in life expectancy, and the focus of this paper, is the fact that it is not uniform. Among industrialized countries, there is marked divergence in life expectancy. And even within industrialized countries there are substantial variations in life expectancy and other measures of health status.

UNEQUAL IMPROVEMENTS

Amid the pattern of general improvement in health, there are some significant irregularities. The first is among countries. *Figure 1* shows life expectancy in 1965 and 1990 for a range of countries. During this period Japan leapfrogged its way to the top of the life expectancy chart.³ Starting from the same position as the United Kingdom (sixty-eight years), life expectancy in Japan improved by eight years, compared to only five years in the United Kingdom.

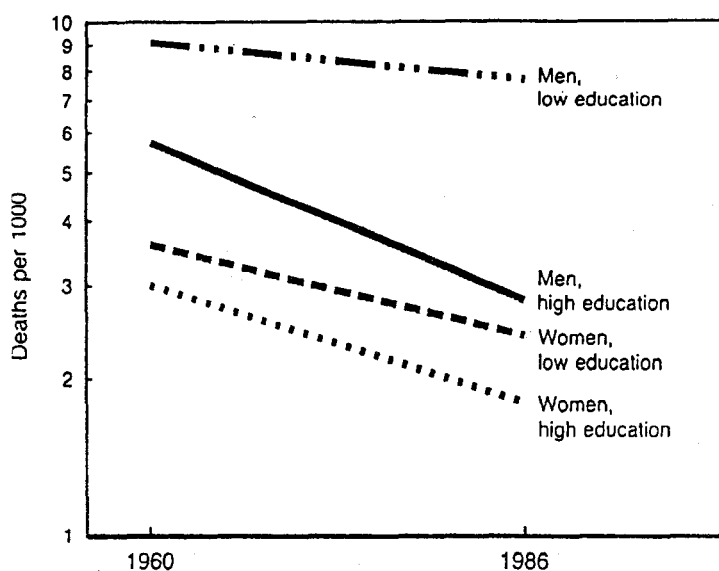
FIGURE 1. Life Expectancy at Birth (Men)



Why did life expectancy in Japan increase so dramatically? Certainly not because of improved behaviors: their dietary fat intake increased during this period, and smoking is prevalent. Furthermore, they spend little on medical care. Could it be that Japan's health success was tied up with its economic success? This might also explain Singapore's improved life expectancy and the lack of improvement in the countries of Central and Eastern Europe. Such differences may relate more to socioeconomic differentials than simply to average prosperity.

The second major irregularity in the pattern of overall improvement is the difference in mortality rates among socioeconomic groups within countries. *Figure 2* compares mortality rates in the United States in 1960 and 1986 according to level of education.⁴ Although mortality rates declined in all cases, the decline was greater for those who received more education. A similar scenario was reported in England and Wales.⁵ In addition, comparisons of mortality rates of those in lower status (manual) and those in higher status (nonmanual) occupations show a widening gap between the two groups.

FIGURE 2. Death Rates in 1960 and 1986 among Whites Twenty-five through Sixty-four Years of Age, according to Sex and Educational Level



This picture of a widening gap in mortality rates between socioeconomic groups at a time of overall improvement has attracted criticism.⁶ It has been argued that, in the United Kingdom, the apparent higher relative mortality of those in lower social classes applies to a dwindling portion of the population. The US analyses allowed for this by comparing relative mortality for constant percentiles of the distribution in the two time periods. The disparity in health status between socioeconomic groups is not disappearing as overall health status of the population improves.

HEALTH, HEALTH CARE, WHO CARES?

When health is the subject of news reports, which is quite often these days, it generally takes one of two forms: titillating snippets that supply a dedicated readership with a wide range of advice, often conflicting, on the life-style route to eternal life; and the funding and organization of health care. The first of these, the life-style advice, may have some impact on the health of a small minority of interested people, but neither is likely to have much impact on the health status of the population as a whole. Thomas McKeown, one of the few to be interested in discovering the determinants of the decades-long improvement in health, argued persuasively that improvements came from nutrition and environmental change, not from advances in medical care.⁷

There is evidence that suggests that extending medical coverage to the whole population is unlikely materially to influence socioeconomic differences in life expectancy.⁸ Why, then, does health-care reform top the political and public agenda in the United States? That nation spends about one-seventh of its national income on medical care. Keeping costs under control will help individuals and companies who have to foot the bill. And extending coverage to those who are presently uninsured will bring the United States more in line with other advanced countries.

If we take life expectancy to be a measure of a country's health status, the United States, despite leading the world in medical-care expenditure, trails rather badly. In 1991, life expectancy for men in the United States was seventy-two years, which was among the worst in the OECD countries.⁹ This pales in comparison to Japan (seventy-six years), Sweden and Hong Kong (seventy-five years),

and Norway and Switzerland (seventy-four years). For women, the United States ranked fourteenth out of twenty-two.

The medical profession, unsurprisingly, is concerned principally with the diagnosis and treatment of disease. There is a commitment to investigating the mechanisms of disease, with a view to improving treatment. To the extent that prevention has been discovered, by and large it follows the traditional medical approach of altering the risk status of individuals. As Geoffrey Rose has shown, unless an effort is made to change the risk status of whole populations, the impact on public health will be limited.¹⁰

Recently, the governments of many countries have set population health targets. In the United Kingdom, *The Health of the Nation* chose five key areas and set targets for the year 2000 and beyond.¹¹ In laying out a strategy for how these targets might be achieved, the National Health Service was seen as central but there was explicit recognition that there are determinants of health that lie outside the health-care system. The contradiction is that unfortunately health ranks high on the agenda only of the Department of Health; other areas of public policy that potentially have a profound influence on health do not usually have health as a policy goal. There is, for example, little explicit recognition that economic and social policies may have a profound influence on the health of the population.

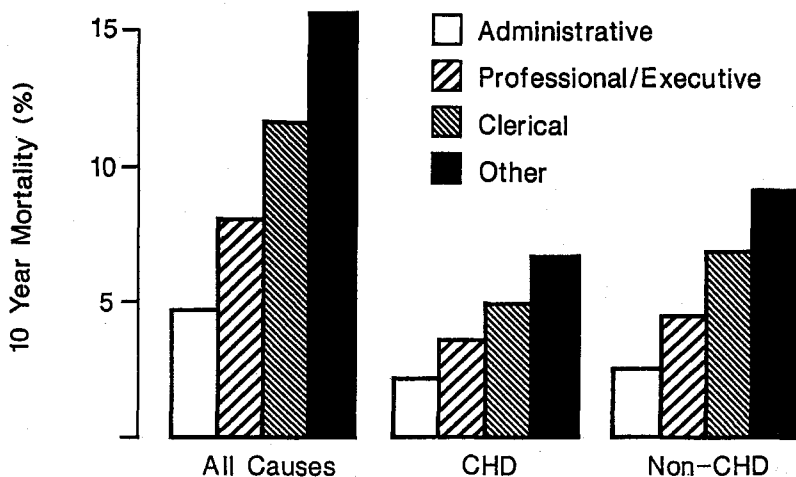
One reason for studying the correlation between socioeconomic differences and health is the insight it provides into social and economic influences on health. Social variations in health status were acknowledged in *The Health of the Nation*, but the conclusion was that the knowledge base was insufficient to recommend specific actions. Taking actions that reduce these differences is not only important as a social goal, but without reductions in socioeconomic differences in health the ability to meet health targets for the whole of society will be reduced.

SOCIOECONOMIC DIFFERENCES WITHIN COUNTRIES: THE WHITEHALL STUDIES

Many studies have shown a relation between various measures of social position and morbidity and mortality.¹² In the original Whitehall study of British civil servants,¹³ we classified men by grade of

employment as an index of social position. The results are shown in *Figure 3*, updated from previous reports.¹⁴

FIGURE 3. Age Adjusted Mortality Rates (percentage) by Grade of Employment for Civil Servants Aged Forty to Sixty-four in the Whitehall Study

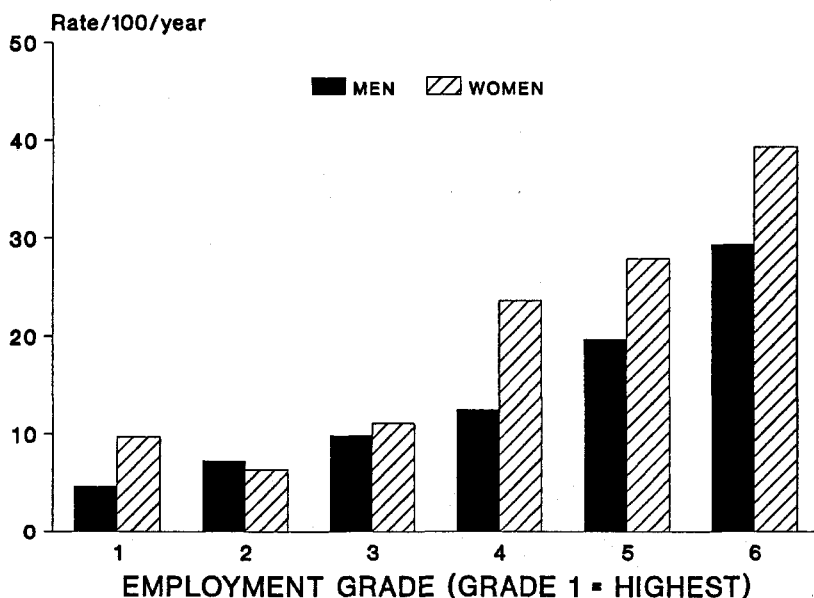


The most striking finding from this study is the social gradient in death rates. The participants in the Whitehall study were predominantly from a single ethnic group; were employed in office jobs, not subject to industrial hazards, unemployment, or extremes of poverty or affluence; and were all working and living in the Greater London area. In this relatively homogeneous population, we observed a gradient in mortality according to job hierarchy—each group had a higher mortality rate than the group one step higher in the hierarchy. The difference in mortality was threefold between the highest and lowest positions in the hierarchy. The question is not why people at the bottom have worse health but why social differentials in health are spread across the whole of society. The task now is to go from description to explanation.

In a second study of civil servants, the Whitehall II study, we have been approaching health as a combination of social, psychological, and physical functioning. As one way of studying global health, we have been examining the determinants of sickness absence, i.e., absence from work due to sickness. *Figure 4* shows sickness absence

rates by grade of employment, with 1 being the highest grade (administrators) and 6 being the lowest grade (clerical officers and office support staff). For both men and women, the rate of absences due to sickness rises sharply from low rates among senior administrators to high rates among lower grade office staff. Here too, what is interesting is not the contrast between those at the top and bottom, but that there is a continuous inverse relation.

FIGURE 4. Long Spells of Sickness Absence



The relation between social position and sickness absence rates holds for short spells of absence (seven days or fewer) as well as for long spells that require a certificate signed by a doctor. If all illness is a mix of biological, psychological, and social dysfunctions, it might be that short spells of absence reflect more of the psychological and social aspects of illness, and long spells more of the biological aspects of illness. Even if short spells reflect psychological and social functioning, not "true" illness, it is interesting that the social gradient in absence rates is identical for short and long spells. One possible reason for this is that the factors responsible for the social gradient in social and psychological functioning are similar to those responsible for the social gradient in "true" illness.

Poverty or Inequality

Much of the discussion about social inequalities in health has focused on the health disadvantage of the lower classes.¹⁵ This is analogous to seeing social problems as particular to a disadvantaged minority, rather than as a problem for society as a whole. There is little doubt that poverty, or deprivation, is likely to be deleterious, not least to health. Young blacks in Harlem have six times the mortality rate of the New York average. The mean income of blacks is less than one-third the US average. The life expectancy of black men in Harlem is below that of men in Bangladesh.¹⁶

The Whitehall data on mortality (*Figure 3*) and Whitehall II data on sickness absence (*Figure 4*) suggest that something other than low levels of income is at work here. This social gradient in mortality suggests the operation of factors across the whole of society. Whether it is relative deprivation or relative lack of access to the fruits of a wealthy society, it is clear that explanations for socio-economic differentials in Britain in the 1990s must be broader than those advanced earlier in the century.¹⁷

This relates to the question with which we began: why has life expectancy improved and how does that relate to the general improvement in living standards? Above a certain level, it is not easy to see how improvements in material well-being would lead to further improvements in health. Is a color television better for health than a black and white one? Is a holiday in Spain better than one on the English coast?

Richard Wilkinson shows that, comparing countries, there is a relation between GNPPc and life expectancy at birth for poor countries: in 1984, few countries with a GNPPc of under \$5,000 had a life expectancy of seventy years or more. Beyond that level, however, there is little relation between GNP and life expectancy. The relation is much closer with measures of income dispersion. Using as a measure of equality of income distribution the share of total household income (after taxes) received by the least well-off 70 percent of families, there is a striking correlation with life expectancy. Wilkinson found that the higher the share of total income enjoyed by the bottom 60 or 70 percent, the longer the average life expectancy of that group. He also showed that changes

in life expectancy correlated with changes in income distribution. Among six countries for which these data were available, Japan had the greatest increase in equality of income distribution and the greatest increase in life expectancy. Japan now has the most equitable distribution of income and the longest life expectancy of any OECD country. By contrast, in the United Kingdom the bottom 60 percent of households had a declining share of total income and the population had a relatively small increase in life expectancy.¹⁸

There are two broad explanations for the relation between income inequality and life expectancy: greater income inequality implies that a greater proportion of the population lives in absolute poverty; or that the determining factors are related to inequality rather than to poverty. If there were a nonlinear relationship between income and life expectancy—stronger at low incomes than at high—the apparent correlation of income inequality and life expectancy could possibly be explained by the differing proportions of those in poverty. Wilkinson points out that this is not the most likely explanation.¹⁹ If the criterion of income inequality is the share of income received by the bottom 10 percent, the relation with life expectancy is very much weaker than if the criterion is the share received by the bottom 60 or 70 percent. Even if increased income inequalities did result in more households in poverty, the proportion of the absolute poor in OECD countries is still too small to account for the size of the relation between income inequality and life expectancy. There must be an effect of income inequality on the large bulk of the population. In other words, the main influence on life expectancy among rich countries is likely to be relative rather than absolute deprivation.

One Disease or Many?

A striking feature of social class differences in health is that the findings apply to diverse pathological conditions. In the Whitehall study, for example, the higher risk of death among those lower in the job hierarchy applied to deaths from lung cancer, other cancers, coronary heart disease (CHD), cerebrovascular disease, other cardiovascular diseases, chronic bronchitis, other respiratory diseases, gastrointestinal disease, genitourinary disease, accidents, and violence.²⁰ These findings have suggested that there may be certain factors that increase general susceptibility to ill health.²¹

An alternative to the general susceptibility hypothesis is that a variety of specific factors are operating to explain social class differences in mortality. Support for this view comes from the fact that some cancers, notably those of the colon, brain, prostate, hematopoietic system, and breast, and melanoma, do not show the same social class variation as the causes listed above.²²

The general susceptibility hypothesis implies that certain groups will be at a higher risk of death whatever the causes. It does not, however, deny the operation of specific causes. For example, diseases caused by smoking, such as chronic bronchitis and lung cancer, show a particularly strong social class gradient. Put another way, according to the general susceptibility hypothesis, there are unknown factors that will increase risk of death from a number of diseases in addition to the known factors such as smoking. This can account for the fact that an administrator who smokes twenty cigarettes a day has a lower risk of dying from lung cancer than does a lower grade civil servant who smokes the same amount,²³ even after pack-years (packs per year times the number of years a person has smoked), tar content, and the gradient in mortality from CHD among nonsmokers are taken into account.²⁴

This increased susceptibility to CHD may well be the result of social factors. People's life courses vary, to a great extent, by social location. And negative effects on health may accumulate throughout a person's life. That influences on health cluster in such a way as to produce among social groups differing degrees of disadvantage with respect to susceptibility to most diseases is undeniable. Unfortunately, our current level of knowledge regarding this general susceptibility does not allow us to go much beyond empirical observation.

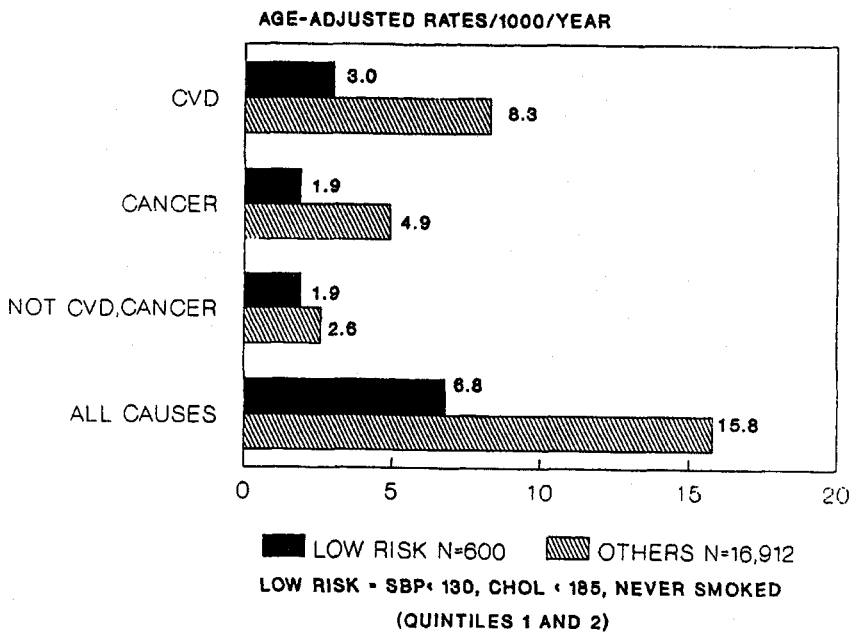
SOCIAL AND OTHER CAUSES OF DISEASE: CHD

Coronary heart disease illustrates why a biological and a social model of the causation of disease must be reconciled. There are two related challenges in the search for causes of the inverse association between socioeconomic position and risk of CHD and other diseases: 1) The social distribution of CHD has changed. In the twentieth century in Britain and elsewhere, CHD, previously more common in higher socioeconomic groups, became more common in lower

socioeconomic groups.²⁵ 2) Much is known about the life-style causes of CHD in addition to the possible role of genetic predisposition.

Using data from the Whitehall study, *Figure 5* shows one approach to explaining why CHD is of epidemic proportions in most industrialized countries. The low risk subgroup was comprised of individuals who had low plasma cholesterol levels (the bottom 40 percent of the population distribution), who had low blood pressure (also the bottom 40 percent), and who had never smoked. If risk factors cause disease, and the evidence suggests that they do, and if it were possible to have a British population in which the "average" level of risk was reduced to that of the low risk group, then rates of mortality from cardiovascular disease would be more than 60 percent lower than at present.

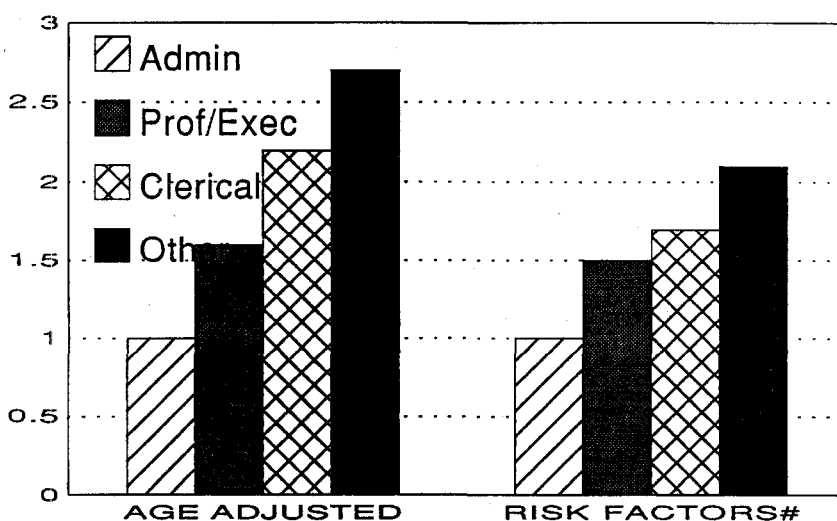
FIGURE 5. Mortality in Whitehall Civil Servants, Men Aged Forty to Sixty-four



If two-thirds of all deaths from cardiovascular disease can be attributed to the three risk factors defined above, how can it also be argued that there are important social determinants? First, *Figure 5*

is a guide to what is biologically feasible, not necessarily practically achievable. Second, given that the British population has *uniformly* high plasma cholesterol levels, cholesterol levels cannot be responsible for the marked social gradient in CHD. In the Whitehall study, mean plasma cholesterol showed a small positive association with grade of employment (higher levels in higher grades); there was a small inverse association with blood pressure; and although the differences in smoking were marked, the social gradient in CHD mortality among nonsmokers was the same as among smokers. *Figure 6* shows that adjusting for coronary risk factors explained about 25 percent of the social gradient in CHD.

FIGURE 6. Relative Risk of CHD Death in Ten Years (controlling for age and other risk# factors)



#age, smoking, blood pressure, cholesterol, height, and blood sugar

Another way of interpreting *Figure 6* is that without a high level of risk factors, CHD would not be a major cause of death for civil servants in Britain. Given the uniformly high level of risk factors in Britain, something must account for the gradient in mortality from CHD, something related to social status that affects susceptibility to CHD. The data show that a number of causes of morbidity and

mortality have a similar social gradient, suggesting that some factors of susceptibility are common to a number of different diseases.

EXPLANATIONS OF SOCIAL INEQUALITIES IN HEALTH

Although social differences in the provision of *medical care* in the United States are pointed to as an explanation of existing social inequalities in health in that country, in the United Kingdom they are a less likely explanation because of universal access under the National Health Service. One way medical care's contribution to health has been assessed is by separating causes of death into two categories: those thought to be amenable to medical intervention and those thought not to be amenable (i.e., where medical care is judged to make no difference to the mortality rate). In the United Kingdom, the improvement in mortality over time is largely the result of the decline in causes of death judged to be nonamenable to medical care. The improvement in mortality has been greater for higher than for lower social classes because there has been a greater decline in the causes of death among the higher classes.

Just as medical care plays a limited role in generating social inequalities in health, it also plays a limited role in generating international inequalities in health. For example, Japan, like the United Kingdom, spends a relatively small proportion of its GNP on medical care. Unlike the United Kingdom, however, Japan experienced a decline in mortality for both amenable and nonamenable causes of death.²⁶

A second possible explanation for social inequalities in health is *health selection*, the idea that health may determine social position rather than the reverse. For example, people with schizophrenia tend to be downwardly mobile. However, there is no evidence to support health selection as an explanation of broader social inequalities in health.²⁷

Factors Operating Early in Life

A different version of the health selection hypothesis, "indirect selection,"²⁸ suggests that common factors determine both social position and health in adulthood. This hypothesis recognizes that people's adult lives are influenced by factors from their earlier days: genetic, biological, educational, cultural, psychological, and social.

It has been convincingly argued that the major influences on health in adulthood operate early in life,²⁹ and it is quite possible that social position is also determined early in life. Yet, the theory regarding the childhood origins of adult disease has been criticized precisely because of the ways in which influences from early life determine the lives people lead and the social environments in which they live and work.³⁰ It may, in fact, be these conditions of adult life that are related to ill health, and the importance of childhood conditions may therefore be indirect.

Whitehall data relating height to mortality help to distinguish between the direct effects on health of early and later life experiences. For example, height, which is both environmentally and genetically influenced, predicts adult mortality independent of grade of employment,³¹ and grade of employment, which is to some extent an index of current social influences, predicts mortality independent of height. Thus, two sets of influences may affect mortality risk: factors from early life and current influences.

Individuals may bring more than biological characteristics from early life to adult life. Education has been shown to be a powerful predictor of health status. This may be because education is a good predictor of social status. People with higher levels of education tend to have better jobs and higher social standing, and the advantages that go along with them. The correlation between education and adult social position may also tell us something about causal pathways. People with more education may have more life management skills, which in turn may predict both adult socioeconomic circumstances and health status.

Health-Related Behaviors and Biological Risk Factors

Why are there links between social position and life-style behaviors such as smoking? There must, of course, be biological pathways underlying the link between social position and disease rates. These biological changes may, in part, result from behaviors such as diet and smoking, but they may also be the result of other influences, such as stress, acting directly on the neuroendocrine system.

Material Conditions

The Black Report emphasized the role of material conditions in causing social inequalities in health.³² Black emphasized the point

that some people have no choice but to be exposed to certain hazards given the present distribution of income and opportunity. This can be interpreted to include psychosocial influences that are inherently related to societal position.

When poor living conditions meant polluted water, crowded, unsanitary housing, and appalling working conditions, it was easy to see how socioeconomic status related to poor health. As countries became wealthier, mortality rates of all social groups improved. But why does the social gradient persist? Are we to understand that there are residual effects of bad housing, as well as air pollution and other material conditions, that, although they affect the lower social groups less than they used to, still affect them more than the higher social strata? If there are such residual effects, it is no surprise that they do not affect all social strata equally. Can this be the whole explanation?

In the Whitehall and Whitehall II studies, morbidity and mortality varied linearly with grade of employment. It is possible, but unlikely, that the second highest grade has worse health than the highest grade because of worse housing, a poorer diet, or more pollution. It also seems unlikely that comfortable "middle class" people in Britain are suffering from the effects of material deprivation.

In fact, what may be important is not absolute deprivation but relative deprivation. Comparing rich countries, Wilkinson showed that it was not differences in wealth that predicted differences in life expectancy, but differences in income distribution. This goes some distance toward explaining the social gradient in health, because each group, while not necessarily deprived, will, by definition, have "less" than the group above it. Thus, the social gradient in health will vary in magnitude depending on the magnitude of the relative differences in "deprivation."

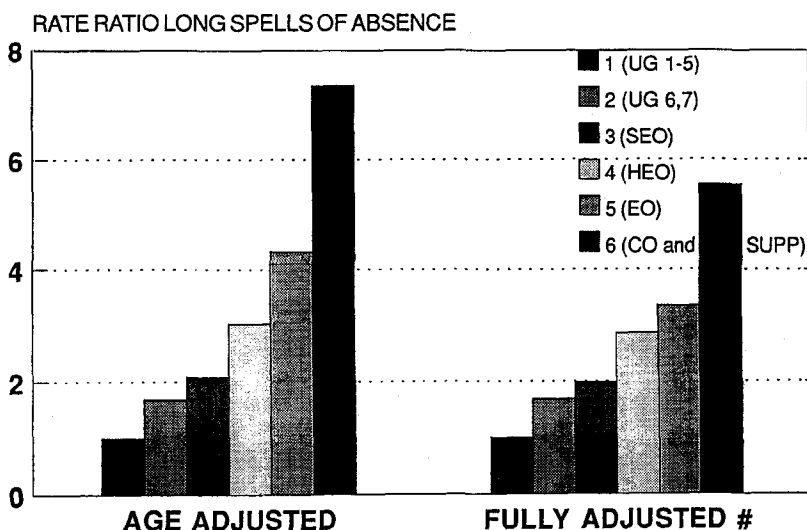
Psychosocial Factors as Potential Explanations

We have established that relative deprivation is an important cause of social differentials in health, but what does it mean and how does it operate? How could a lower position in the social hierarchy cause heart disease and cancer, among other diseases, if not by diet or smoking?

A number of psychosocial factors have been shown to be related to mortality and morbidity,³³ among them: less sense of control over one's work; lower use of skills and less variety on the job; lower participation in social networks and activities; fewer social supports; and less sense of personal control over health. Data from the Whitehall II study show that lower position in the occupational hierarchy is associated with disadvantage on all of these measures.³⁴ Syme has suggested that increasing lack of control may account for increasing health disadvantage down the social scale.³⁵ Relative poverty, whatever else it may represent, leads to a lack of control.

Figure 4 showed a steep gradient in sickness absence rates according to employment grade. Three sets of characteristics were related to both grade of employment and risk of sickness absence: personal behaviors, such as smoking; work environment; and social circumstances outside work.³⁶ Figure 7 shows sickness absence rates by grade. Allowing only for age in the first instance and then for the other effects on sickness absence helps us determine the extent to which these characteristics account for the grade/sickness absence relationship. In fact, they are responsible for about a third.

FIGURE 7. Long (>7D) Absence by Grade, Men Whitehall II Study



#effects of smoking, psychosocial work environment, psychiatric symptoms, social environment outside work (financial difficulties and negative aspects of social support), ethnic group

*Grade ranking is from 1=senior administrators (high) to 6=clerical officers and office support staff (low)

At first glance, this is somewhat disappointing. Why can we not explain more than a third of the social gradient in sickness absence rates? Measurement issues must be taken into account. Peto and colleagues showed that measurement imprecision leads to a 50 percent underestimate of the relation between blood pressure level and cardiovascular disease.³⁷ There is a similar underestimate for cholesterol level. It may also be the case that measurement imprecision has underestimated the influence of these psychosocial factors. Second, little effort has been given to understanding the interrelation of these psychosocial factors and the different meaning they may have at different points in the hierarchy. For example, we have the impression that having to work fast is a positive characteristic for people in high grades, but it may represent psychological demands for people in low grades. Third, there may be some crucial variables that have been overlooked.

CONCLUSION

Improvements in health have become expected in wealthy countries. But both within and between countries there are substantial inequalities in health that have remained despite overall improvements. In some cases, these inequalities have become even more clearly visible.

While universal access to high quality, affordable medical care is a social goal that should have high priority, it alone will not eliminate social inequalities in mortality rates. The fact that inequalities in health are not limited to worse health among the poor but that a gradient runs across society, in combination with the close link between income inequalities and health, indicates that relative deprivation is an important determinant of health inequalities. Such deprivation influences life-style choices and differential access to high quality social environments. Providing more people with fulfilling jobs, adequate compensation, and social environments that foster good relationships may be of crucial importance in reducing inequalities in health.

ENDNOTES

¹Leonard A. Sagan, *The Health of Nations, True Causes of Sickness and Well-Being* (New York: Basic Books, 1987), 1-233.

- ²Richard G. Wilkinson, "Income Distribution and Life Expectancy," *British Medical Journal* 304 (1992): 165–68.
- ³Michael G. Marmot and George Davey-Smith, "Why are the Japanese Living Longer?," *British Medical Journal* 299 (1989): 1547–551.
- ⁴Gregory Pappas, Susan Queen, Wilbur Hadden, and Gail Fisher, "The Increasing Disparity in Mortality between Socioeconomic Groups in the United States, 1960 and 1986," *New England Journal of Medicine* 329 (1993): 103–109.
- ⁵Michael G. Marmot and Michael E. McDowall, "Mortality Decline and Widening Social Inequalities," *Lancet* I (1986): 274–76.
- ⁶Raymond Illsley, "Occupational Class, Selection and the Production of Inequalities in Health," *Quarterly Journal of Social Affairs* 2 (1986): 151–65. Philip M. Strong, "Black on Class and Mortality: Theory, Method and History," *Journal of Public Health Medicine* 12 (1990): 168–80.
- ⁷Thomas McKeown, *The Role of Medicine* (London: Nuffield Provincial Hospitals Trust, 1976).
- ⁸Michael G. Marmot, Martin Bobak, and George Davey-Smith, "Explanations for Social Inequalities in Health," in Sol Levine, Benjamin Amik, Alvin R. Tarlov, and Diana Chapman Walsh, eds., *Society and Health* (forthcoming).
- ⁹World Bank, *World Development Report*, 1992 (Oxford: Oxford University Press, 1992).
- ¹⁰Geoffrey Arthur Rose, *The Strategy of Preventive Medicine* (Oxford: Oxford University Press, 1992).
- ¹¹HMSO, *The Health of the Nation* (London: HMSO, 1992).
- ¹²Michael G. Marmot, Manolis Kogevinas, and Mary Anne Elston, "Social/Economic Status and Disease," *Annual Review of Public Health* 8 (1987): 111–37.
- ¹³Michael G. Marmot, Geoffrey Rose, Martin Shipley, and P. J. S. Hamilton, "Employment Grade and Coronary Heart Disease in British Civil Servants," *Journal of Epidemiology and Community Health* 32 (1978): 244–49.
- ¹⁴Michael G. Marmot, Martin J. Shipley, and Geoffrey Rose, "Inequalities in Death—Specific Explanations of a General Pattern?," *Lancet* I (1984): 1003.
- ¹⁵George C. M. M'Gonigle and John Kirby, *Poverty and Public Health* (London: Gollancz, 1937).
- ¹⁶Colin McCord, "Excess Mortality in Harlem," *New England Journal of Medicine* 322 (1990): 173–77.
- ¹⁷M'Gonigle and Kirby, *Poverty and Public Health*.
- ¹⁸Wilkinson, "Income Distribution and Life Expectancy."
- ¹⁹Ibid.
- ²⁰Marmot, Shipley, and Rose, "Inequalities in Death—Specific Explanations of a General Pattern?."
- ²¹Lisa F. Berkman and S. Leonard Syme, "Social Networks, Host Resistance and Mortality: A Nine-Year Follow-up of Alameda County Residents," *American*

- Journal of Epidemiology* 109 (1979): 186–204. John C. Cassel, "The Contribution of the Social Environment to Host Resistance," *American Journal of Epidemiology* 104 (1976): 107–23.
- ²²George Davey-Smith, David Leon, Martin J. Shipley, and Geoffrey Rose, "Socio-economic Differentials in Cancer Among Men," *International Journal of Epidemiology* 20 (1991): 339–45.
- ²³Marmot, Shipley, and Rose, "Inequalities in Death—Specific Explanations of a General Pattern?."
- ²⁴George Davey-Smith and Martin J. Shipley, "Confounding of Occupation and Smoking: Its Magnitude and Consequences," *Social Science and Medicine* 32 (1991): 1297–300.
- ²⁵Michael G. Marmot, Abraham M. Adelstein, Nicola N. Robinson, and Geoffrey A. Rose, "Changing Social Class Distribution of Heart Disease," *British Medical Journal* 2 (1978): 1109–112.
- ²⁶Marmot and Davey-Smith, "Why are the Japanese Living Longer?."
- ²⁷This was one of the possible explanations considered by *The Black Report* and rejected. Douglas Black, Jeremy N. Morris, Cyril Smith, Peter Townsend, and Margaret Whitehead, *Inequalities in Health: The Black Report* (London: Penguin Group, 1988).
- ²⁸Richard G. Wilkinson, "Occupational Class, Selection and Inequalities in Health: A Reply to Raymond Illsley," *Quarterly Journal of Social Affairs* 2 (1986): 415–22.
- ²⁹David J. P. Barker, "The Intrauterine and Early Postnatal Origins of Cardiovascular Disease and Chronic Bronchitis," *Journal of Epidemiology and Community Health* 43 (1989): 237–40.
- ³⁰Yoav Ben-Shlomo and George Davey-Smith, "Deprivation in Infancy or Adult Life: Which is More Important for Mortality Risk?," *Lancet* 337 (1991): 530–43.
- ³¹Marmot, Shipley, and Rose, "Inequalities in Death—Specific Explanations of a General Pattern?."
- ³²Black, Morris, Smith, Townsend, and Whitehead, *Inequalities in Health: The Black Report*.
- ³³Robert Karasek and Tores Theorell, *Healthy Work: Stress, Productivity, and the Reconstruction of Working Life* (New York: Basic Books, 1990). Michael Marmot and Tores Theorell, "Social Class and Cardiovascular Disease: The Contribution of Work," *International Journal of Health Services* 18 (1988): 659–74. Lisa F. Berkman, "Assessing the Physical Health Effects of Social Networks and Social Support," in Lester Breslow, Jonathan E. Fielding, and Lester B. Lave, eds., *Annual Review of Public Health*, 5th ed. (Palo Alto, Calif.: Annual Reviews Inc., 1984), 413–32. James S. House, Karl R. Landis, and Debra Umberson, "Social Relationships and Health," *Science* 241 (1988): 540–45.
- ³⁴Michael G. Marmot, George Davey-Smith, Stephen Stansfeld et al., "Health Inequalities Among British Civil Servants: The Whitehall II Study," *Lancet* 337 (1991): 1387–393.

- ³⁵S. Leonard Syme, "Control and Health: A Personal Perspective," in Andrew Steptoe and Ad Appels, eds., *Stress, Personal Control and Health* (New York: Wiley & Sons, 1989), 3–18.
- ³⁶Fiona North, S. Leonard Syme, Amanda Feeney et al., "Explaining Socioeconomic Differences in Sickness Absence: The Whitehall II Study," *British Medical Journal* 306 (1993): 361–66.
- ³⁷Stephen MacMahon, Richard Peto, Jeffrey Cutler et al., "Blood Pressure, Stroke, and Coronary Heart Disease. Part 1, Prolonged Differences in Blood Pressure: Prospective Observational Studies Corrected for the Regression Dilution Bias," *Lancet* 335 (1992): 765–74.

DÆDALUS

JOURNAL OF THE AMERICAN ACADEMY OF ARTS AND SCIENCES

Volume 123 of the *Proceedings of the American Academy of Arts and Sciences*, 1994

Contents

Index of Authors



Contents

Germany in Transition

Number 1, Winter 1994

- V Preface
Michael Mertes
- 1 Germany's Social and Political Culture: Change Through
Consensus?
Steven Muller
- 33 Democracy in Germany
Kurt J. Lauk
- 57 Germany at the Crossroads: On the Efficiency of the
German Economy
Klaus J. Bade
- 85 Immigration and Social Peace in United Germany
Heinrich August Winkler
- 107 Rebuilding of a Nation: The Germans Before and After
Unification
Anne-Marie Le Gloannec
- 129 On German Identity
Stephan Eisel
- 149 The Politics of a United Germany
Jürgen Kocka
- 173 Crisis of Unification: How Germany Changes
Ludger Kühnhardt
- 193 Multi-German Germany
Mary Fulbrook
- 211 Aspects of Society and Identity in the New Germany
Barthold C. Witte
- 235 Two Catastrophes, Two Causes, and How the Germans
Dealt with Them

- Richard Schröder**
251 The Role of the Protestant Church in German Unification
- Jochen Thies**
263 Observations on the Political Class in Germany
- Joachim Gauck**
277 Dealing with a Stasi Past

Contents

Europe Through a Glass Darkly

Number 2, Spring 1994

- V Preface
Stanley Hoffmann
- 1 Europe's Identity Crisis Revisited
Philip R. Schlesinger
- 25 Europe's Contradictory Communicative Space
Fabio Luca Cavazza and Carlo Pelanda
- 53 Maastricht: Before, During, After
Hans-Peter Schwarz
- 81 Germany's National and European Interests
Luigi Campiglio
- 107 Europe on the Mark: Ready to Go?
Mark Lilla
- 129 The Other Velvet Revolution: Continental Liberalism and
its Discontents
Nancy Bermeo
- 159 Democracy in Europe
Jack Snyder
- 179 Russian Backwardness and the Future of Europe
Stevan K. Pavlowitch
- 203 Who is "Balkanizing" Whom? The Misunderstandings
Between the Debris of Yugoslavia and an Unprepared West

Contents

After Communism: What?

Number 3, Summer 1994

- V Preface
Tony Judt
- 1 Nineteen Eighty-Nine: The End of *Which* European Era?
Roman Szporluk
- 21 After Empire: What?
Richard Rose
- 41 Getting By Without Government: Everyday Life in Russia
Mark Medish
- 63 Russia: Lost and Found
Jacques Rupnik
- 91 Europe's New Frontiers: Remapping Europe
Elemér Hankiss
- 115 European Paradigms: East and West, 1945–1994
George Schöpflin
- 127 Postcommunism: The Problems of Democratic Construction
George Kolankiewicz
- 143 Elites in Search of a Political Formula
István Rév
- 159 The Postmortem Victory of Communism
Andrei Marga
- 171 The Culture of Scholarship in Europe Today

NOTES FROM THE PAST

- Eric Weil
- 185 Philosophical and Political Thought in Europe Today

Contents

Health and Wealth

Number 4, Fall 1994

V Preface

John W. Frank and J. Fraser Mustard

1 The Determinants of Health from a Historical Perspective

Robert G. Evans

21 Health Care as a Threat to Health: Defense, Opulence, and the Social Environment

John N. Lavis and Gregory L. Stoddart

43 Can We Have Too Much Health Care?

Richard G. Wilkinson

61 The Epidemiological Transition: From Material Scarcity to Social Disadvantage?

S. Leonard Syme

79 The Social Environment and Health

Naoko T. Miyaji and Margaret Lock

87 Monitoring Motherhood: Sociocultural and Historical Aspects of Maternal and Child Health in Japan

Naoki Ikegami

113 Efficiency and Effectiveness in Health Care

Nathan Rosenberg

127 How the Developed Countries Became Rich

Paul M. Romer

141 Economic Growth and Investment in Children

Max S. Cynader

155 Mechanisms of Brain Development and Their Role in Health and Well-Being

Clyde Hertzman

167 The Lifelong Impact of Childhood Experiences: A Population Health Perspective

- Michael C. Wolfson**
181 Toward a System of Health Statistics
- Michael G. Marmot**
197 Social Differentials in Health Within and Between Populations

Index of Authors

- Klaus J. Bade, 85 (*Winter*)
Nancy Bermeo, 159 (*Spring*)
Luigi Campiglio, 107 (*Spring*)
Fabio Luca Cavazza, 53 (*Spring*)
Max S. Cynader, 155 (*Fall*)
Stephan Eisel, 149 (*Winter*)
Robert G. Evans, 21 (*Fall*)
John W. Frank, 1 (*Fall*)
Mary Fulbrook, 211 (*Winter*)
Joachim Gauck, 277 (*Winter*)
Elemér Hankiss, 115 (*Summer*)
Clyde Hertzman, 167 (*Fall*)
Stanley Hoffmann, 1 (*Spring*)
Naoki Ikegami, 113 (*Fall*)
Tony Judt, 1 (*Summer*)
Jürgen Kocka, 173 (*Winter*)
George Kolankiewicz, 143 (*Summer*)
Ludger Kühnhardt, 193 (*Winter*)
Kurt J. Lauk, 57 (*Winter*)
John N. Lavis, 43 (*Fall*)
Anne-Marie Le Gloannec, 129 (*Winter*)
Mark Lilla, 129 (*Spring*)
Margaret Lock, 87 (*Fall*)
Andrei Marga, 171 (*Summer*)
Michael G. Marmot, 197 (*Fall*)
Mark Medish, 63 (*Summer*)
Michael Mertes, 1 (*Winter*)
Naoko T. Miyaji, 87 (*Fall*)
Steven Muller, 33 (*Winter*)
J. Fraser Mustard, 1 (*Fall*)
Stevan K. Pavlowitch, 203 (*Spring*)
Carlo Pelanda, 53 (*Spring*)
István Rév, 159 (*Summer*)
Paul M. Romer, 141 (*Fall*)
Richard Rose, 41 (*Summer*)
Nathan Rosenberg, 127 (*Fall*)
Jacques Rupnik, 91 (*Summer*)
Philip R. Schlesinger, 25 (*Spring*)
George Schöpflin, 127 (*Summer*)
Richard Schröder, 251 (*Winter*)
Hans-Peter Schwarz, 81 (*Spring*)
Jack Snyder, 179 (*Spring*)
Gregory L. Stoddart, 43 (*Fall*)
S. Leonard Syme, 79 (*Fall*)
Roman Szporluk, 21 (*Summer*)
Jochen Thies, 263 (*Winter*)
Eric Weil, 185 (*Summer*)
Richard G. Wilkinson, 61 (*Fall*)
Heinrich August Winkler, 107 (*Winter*)
Barthold C. Witte, 235 (*Winter*)
Michael C. Wolfson, 181 (*Fall*)

The shifts in emphasis from illness to health, from therapeutic to preventive medicine, and from the dominance and autonomy of the doctor to patient's rights and greater control of the medical profession do not alter the fact that health, illness, and medicine are central preoccupations in the society which have diffuse symbolic as well as practical meaning. All signs suggest that they will maintain the social, ethical, and existential significance they have acquired, even though by the year 2000 some structural aspects of the way that medicine and care are organized and delivered may have changed. In fact, if the issues now being considered under the rubric of bioethics are predictive of what lies ahead, we can expect that in the future, health, illness, and medicine will acquire even greater importance as one of the primary symbolic media through which American society will grapple with fundamental questions of value and belief. What social mechanisms we will develop to come to terms with these "collective conscience" issues, and exactly what role physicians, health professionals, biologists, jurists, politicians, philosophers, theologians, social scientists, and the public at large will play in their resolution remains to be seen. But it is a distinctive characteristic of an advanced modern society like our own that scientific, technical, clinical, social, ethical, and religious concerns should be joined in this way.

Renée C. Fox

From "The Medicalization and Demedicalization
of American Society"
Dædalus 106 (1) (Winter 1977)

***DÆDALUS* Issues in Print**

After Communism: What? (Summer 1994)
Europe Through a Glass Darkly (Spring 1994)
Germany in Transition (Winter 1994)
Reconstructing Nations and States (Summer 1993)
America's Childhood (Winter 1993)
Immobile Democracy? (Fall 1992)
Political Pharmacology: Thinking About Drugs (Summer 1992)
The Exit from Communism (Spring 1992)
Searching for Security in a Global Economy (Fall 1991)
Religion and Politics (Summer 1991)
Arms Control: Thirty Years On (Winter 1991)
Risk (Fall 1990)
Literacy in America (Spring 1990)
Another India (Fall 1989)
Three Decades of *DÆDALUS* (Summer 1988)
Religion and Education (Spring 1988)
Learning About Women: Gender, Politics, and Power (Fall 1987)
Futures (Summer 1987)
Past and Present (Spring 1987)

* * *

To order send \$10.95 per issue (includes all shipping and handling)
to: *Dædalus* Business Office • 136 Irving Street • Cambridge, MA
02138

*For orders outside the United States, add \$1.00 per volume for
surface mail delivery.*

How to Get Permission to Reprint *Dædalus* Articles

If you wish to reprint an article from *Dædalus* in another publication or to reproduce an article for classroom or other use, please send a written request to:

Permissions Manager
Dædalus
136 Irving Street
Cambridge, MA 02138

Permission can be granted in most cases; charges vary according to use of the copyrighted materials. If you have any questions, call (617) 491-2600 between 9:00 A.M. and 5:00 P.M. Eastern Time, and ask to speak with our permissions manager.

John W. Frank and J. Fraser Mustard
The Determinants of Health from a Historical Perspective

Robert G. Evans
Health Care as a Threat to Health: Defense, Opulence,
and the Social Environment

John N. Lavis and Gregory L. Stoddart
Can We Have Too Much Health Care?

Richard G. Wilkinson
The Epidemiological Transition: From Material Scarcity
to Social Disadvantage?

S. Leonard Syme
The Social Environment and Health

Naoko T. Miyaji and Margaret Lock
Monitoring Motherhood: Sociocultural and Historical
Aspects of Maternal and Child Health in Japan

Naoki Ikegami
Efficiency and Effectiveness in Health Care

Nathan Rosenberg
How the Developed Countries Became Rich

Paul M. Romer
Economic Growth and Investment in Children

Max S. Cynader
Mechanisms of Brain Development and Their Role in
Health and Well-Being

Clyde Hertzman
The Lifelong Impact of Childhood Experiences: A
Population Health Perspective

Michael C. Wolfson
Toward a System of Health Statistics

Michael G. Marmot
Social Differentials in Health Within and Between
Populations

